DOCUMENT RESUME

ED 096 822 FL 006 216

AUTHOR Harries, Helga
TITLE Contrastive Emphasis and Cleft Sentences. Working Papers on Language Universals, No. 12.
INSTITUTION Stanford Univ., Calif. Committee on Linguistics.
PUB DATE Nov 73
NOTE 60p.

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS *Contrastive Linguistics; *Deep Structure; Grammar; *Language Universals; Linguistic Theory; *Sentence Structure; Surface Structure; Synchronic Linguistics; *Syntax; Transformation Generative Grammar

ABSTRACT

The purpose of this paper is to investigate how languages express contrastive emphasis. It is argued that all contrastively emphasized constructions have underlying cleft sentences, independent of whether the surface structure is an equational or a nonequational one. It is furthermore argued that emphatic word orders are systematic and predictable given a certain language type, and that the position of the object plays an essential role both in the cleft and noncleft emphatic constructions. (Author)

ED 096822

Working Papers on Language Universals
No. 12, November 1973
pp. 85-144

CONTRASTIVE EMPHASIS AND CLEFT SENTENCES

Helga Harries
Language Universals Project

ABSTRACT

The purpose of this paper is to investigate how languages express contrastive emphasis. It is argued that all contrastively-emphasized constructions have underlying cleft sentences, independent of whether the surface structure is an equational or a non-equational one. It is furthermore argued that emphatic word orders are systematic and predictable given a certain language type, and that the position of the object plays an essential role both in cleft and non-cleft emphatic constructions.

FL 006216

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Helga Harries
Delisle
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

1. Introduction

Contrastive emphasis is used by a speaker to mark a constituent as being in contrast with another structurally identical constituent.[1] The contrast made can be explicit as in (1) or implicit as in (2):[2]

(1) a. John bought a camel.
b. No, John didn't buy a camel, John bought a donkey.

(2) a. John bought a camel.
b. No, John bought a donkey.

Languages use syntactic as well as phonological means to express contrastive emphasis: the phonological means are increased loudness (stress) and a concomitant rise in pitch, and tonal changes, while the syntactic means comprise emphatic word orders, emphatic morphemes, and an emphatic sentence type, i.e. cleft sentences. It appears that phonological means are secondary, that is there are no languages that solely make use of phonological means to express contrastive emphasis. On the other hand, there are languages that solely use syntactic means to express emphasis. In this paper I will argue for the hypothesis that all sentences that contain a contrastively-stressed constituent are derived from an underlying cleft construction. To these cleft sentences rules apply -- optionally or obligatorily -- reducing the cleft construction to a greater or lesser degree. The ultimate reduction is that of a complex cleft sentence to a surface simplex. The degree of reduction possible differs from language to language.

The above hypothesis is supported by semantic as well as by syntactic evidence. The semantic evidence is the same for all languages, namely sentences as in (3) below have the same presupposition and make the same assertion.

(3) a. The one who bought the camel was Bill.
b. Bill was the one who bought the camel.
c. It was Bill who bought the camel.
d. Bill bought the camel.

On the other hand, the syntactic support for the above hypothesis is extremely diverse, including case marking, negation, and subject-verb

---

[1] In this paper I will limit myself to the discussion of the contrastive emphasis of noun phrases. It is obvious that not only noun phrases but practically any element can be contrastively emphasized. One of the reasons for the above limitation is that grammars hardly ever provide any information about emphasis on non-nominal constituents. Another reason is that contrastive emphasis of other constituent types, for instance of verbs, seems to involve a slightly different mechanism.

[2] Underlining of a constituent indicates that it carries contrastive emphasis.

agreement. However, before considering this evidence, I will briefly investigate the nature of a cleft sentence so that we have certain criteria which will help us to decide whether or not a given sentence can be classified as a cleft.

2. Manifestation of cleft sentences

2.1 Elements of a cleft sentence

The term cleft sentences is used in this paper to refer to sentences like (4.a) to (4.c) below.

(4) a. The one who helped us was Frank.
b. Frank was the one who helped us.
c. It was Frank who helped us.

In recent linguistic writings[3] on the subject, sentences like (4.a) and (4.b) are referred to as pseudo-cleft sentences, while sentences like (4.c) are generally called cleft sentences. In my opinion this distinction is not a deep structure distinction, but is due to a movement rule. Evidence for this hypothesis will be presented later. Cleft sentences as in (4) are equational sentences which establish an identity between a known or presupposed entity and a focussed entity which represents the new information. The presupposed information is contained in the subject, the new information in the predicate. The subject of a cleft sentence consists of a head noun like the one[4] which is modified by a restrictive relative clause. The head noun is always a neutral noun like the one, the man, the person, the he, which is more closely defined by the relative clause.[5] The predicate contains the focus constituent which in the above example is Frank.

The subject-head and modifying relative clause on the one hand and the focus constituent on the other stand in an X = Y identity relation to each other. In some languages this relation is expressed by a copula morpheme, in others by simple juxtaposition. In languages that distinguish between a 'be' verb of existence and the copula, it is the latter that will show up in a cleft construction as for instance in MANDARIN CHINESE.

---

[3] Jespersen, 1949; Lees, 1963; Akmajian, 1970; and Schachter, 1973.

[4] This head noun is only present in (4.a) and (4.b). It was deleted in (4.c).

[5] I will not investigate here what forms restrictive relative clauses can take. For some discussion see Schwartz, 1971, and for a typology of restrictive relative clauses see Keenan, 1972.

(5) copula: mao shìh dùng-wù
cat be animal
'The cat is an animal.'

(6) 'be' of existence: ta dzài wu lǐ
he be room in
'He is in the room.'

(7) dzúo-tien méi lái de rén {*dzài / shìh} Yohàn
yesterday not came RM[6] person be John
'The one who didn't come yesterday was John.'

In some languages cleft constructions can contain all of the above-specified features, as for instance in ENGLISH, GERMAN, MANDARIN CHINESE, JAPANESE, HUNGARIAN, KIHUNG'AN and SWAHILI.

(8) GERMAN:
Derjenige, der segelt, ist mein Bruder.
'The one who is sailing is my brother.'

(9) MANDARIN CHINESE:
wǒ kàn-jiàn de rén shìh Yòhan
I saw RM person be John
'The one who I saw is John.'

(10) JAPANESE:
Mary o butta hito wa Bill da
Mary obj.mk. hit person subj.mk. Bill is.
'The one who hit Mary is Bill.'

(11) HUNGARIAN:
John volt az, aki New York-ba repült.
John was that, who New York-to flew
'John was the one who flew to New York.'

(12) KIHUNG'AN:
kiim ki a - swiim - in Kipes zoon kwe kít
thing pron. PA-buy-past Kipes yesterday is chair
'What Kipes bought yesterday is a chair.

(13) SWAHILI:
mtu huyu ndiye ninayem taka
man this is-he I-am-who-him wanting
'This man is the one I wanted.'

In other languages the surface structure of an underlying clefted construction is more reduced. However, even if certain features of a

---

[6] RM stands for relative marker, which may be a relative pronoun, an invariable linker-subordinator, a nominalizer, etc.

cleft sentence are not present in the surface form of certain languages, other features still allow us to determine whether or not a given sentence is a cleft construction.

2.2 Absence of the copula

It is a well-known fact[7] that many languages do not make use of a copula morpheme to establish an identity between two constituents. In general these languages simply juxtapose the two constituents,[8] e.g.:

(14) TAGALOG:
mataas ang babae
tall the woman
'The woman is tall.'

(15) ARABIC:
Ali mu?alimun.
Ali teacher
'Ali is a teacher.'

(16) JAVANESE:
aku muréd
I student
'I am a student.'

Given that many languages do not make use of a copula in a sentence like John is my brother, we would expect that in these languages the copula can also be absent in a more complex but basically identical construction like The one who came is my brother. This is in fact the case.

(17) TAGALOG:
si Rosa ang siyang maganda
Construct.m. Rosa nomin. he-RM pretty
'Rosa is the one who is pretty.'

(18) ARABIC (COLLOQUIAL):
Ali huwa-lli jā
Ali he -RM came
'Ali is the one who came.'

(19) INDONESIAN:
orang yang kepada siapa saya memberihan buku ini
man RM to resp.pron. I give book this
'The man is the one to whom I gave this book.'

---

[7] See Benveniste, 1960; Bach, 1967.

[8] The identity relation may be signaled by a pause between the two constituents.

2.3 Absence of the neutral head noun

Besides the copula, the neutral head noun can be absent in many languages in a cleft construction.

(20) GERMAN:[9]
Der segelt, das ist mein Bruder.
who is sailing, that is my brother
'The one who is sailing is my brother.'

(21) ENGLISH:
What Frank lost was his watch.

(22) MANDARIN CHINESE:
Yohàn kàn-jièn de shìh ge nán rén
John saw RM be class. male person
'The one who John saw is a man.' (as opposed to a woman)

(23) AMHARIC:
əssu naw yamattaw
he is RM-came
'He is the one who came.'

2.4 Absence of relative clause marker

Another constituent that may or may not occur is the morpheme that marks relative clauses. In ENGLISH its occurrence is optional in some environments.

(24) It is John I saw yesterday.

In DERA, a CHADIC language, in RUSSIAN and in INDONESIAN, both the copula and the relative clause marker can be absent.

(25) RUSSIAN:
Eto ego ja vstretil včera.
it/this he I met yesterday
'It is he that I met yesterday.'

(26) DERA:
wuni wun kapa kurei
they ones plant corn
'They are the ones who plant corn.'

(27) INDONESIAN:
bukan saya yang beladjar bahasa Indónésia
not I RM studying language Indonesian
'I am not the one who studies Indonesian.'

---

[9] In GERMAN head-deletion is for most speakers more acceptable if dislocation of the subject NP has taken place as in the sentence above, das 'that' is the neutral demonstrative pronoun.

2.5 Identical constraints on relative clauses and emphatic sentences

While in some languages the underlying cleft structure of emphatic constructions is revealed by the presence of a copula, a neutral head, and/or a relative clause marker, other languages reveal the presence of an underlying cleft construction by the fact that in these languages relative clauses and emphatic constructions are subject to identical constraints. These constraints can be accounted for if we assume that emphatic constructions are derived from underlying cleft sentences that contain a relative clause. For instance, a number of languages have a special tense form that occurs in relative clauses, emphatic constructions, and usually also in word questions. Sapir (1965) reports for DIOLA-FOGNY, a language of the Senegal, that the verbal of a relative clause is marked by the suffix - ɛ.

(28) balba:b bagi lɛnɛ
sun-the classifier-rel. linker-is cool-
'the cool sun' (i.e. the sun which is cool)

The same marker also shows up in emphatic constructions, where it emphasizes the constituent that precedes the verb.

(29) kukila kujɛ dakar
they went+ɛ Dakar
'They are the ones who went to Dakar.'

(30) dakar kujɛ
Dakar they-went+ɛ
'It is to Dakar that they went.'

In TELEGU, a DRAVIDIAN language, the verb form of an emphatic construction is the relative participle plus the suffix di:

(31) a. neutral: raamaa raawu jarmaniininci kotta kaaru teppincEEDu
Rama Rao from-Germany new car get
'Rama Rao got a new car from Germany.'

b. emphatic: jarmaniininci kotta kaaru teppineindi raamaa raawu
from-Germany new car get Rama Rao
'It is from Germany that Rama Rao got a new car.'

In KIKUYU, the copula nĩ occurs sentence initially in an emphatic construction and the form of the verb is that of a verb of a dependent clause:

(32) a. neutral: nĩn gahaica kĩrĩma kĩu rũciũ
I shall-climb that hill tomorrow
'I shall climb that hill tomorrow.'

b. emphatic: nĩ kĩrĩma kĩu ngahaica rũciũ
is that hill I-shall-climb tomorrow
'It is that hill that I shall climb tomorrow.'

MALAGASY, according to Malzac, has a relative verb form for which he does not explicitly states that it also occurs in emphatic statements. However, with one or two exceptions, all of his examples of the relative verb form involve emphatic constructions containing the emphatic marker no instead of the relative marker izay:

(33) ela no tsy nahitako anao
long-time EM not see-I you
'It is a long time that I didn't see you.'

In KANURI, a SAHARAN language, subject emphasis is expressed by a construction in which the verb, if positive and past tense, is identical with the verb form of the relative past, i.e. the past tense form used in relative clauses. The same relative past also occurs in information questions. In HAUSA, according to Schachter (1973), certain person and aspect markers have to occur instead. Furthermore, certain pronominalization and deletion processes that are involved in the formation of emphasized constructions are identical to the pronominalization and deletion processes in relative clauses.

(34) HAUSA:
sun gaya wa yaron
they told to child

(35) yaron da suka gaya {masa / wa}
child that they told to

(36) yaron ne suka gaya {masa / wa}
child is they told to
'It's the child that they told.'

According to Newman (1970), it is very common in CHADIC languages that relative clauses share syntactic characteristics with emphatic constructions and also with information questions. For instance, in TERA the verb of a neutral declarative is marked by the affix wa in the perfective tense. This affix is absent in relative clauses, emphatic constructions and information questions.

Constraints shared by relative clauses and emphatic constructions are not only restricted to tense and aspect but can also involve negation, agreement, constraints on deletion, and nominalization. Takizala states that in KIHUNG'AN, a BANTU language, relative clauses, emphatic constructions and information questions all use a negative particle which differs from the negative particle of neutral, unembedded declaratives:

(37) lo i-mween kit
neg I-saw chair
'I didn't see the chair.'

(38) kit ki a-khoon-in Kipes ku-suum
chair pron. PA-fail-past Kipes to-buy
'The chair that Kipes didn't buy.'

*kit ki lo a-swiim-in Kipes

(39) Kipes ka-khoon-in ku-suum kit zoono
Kipes PA-fail-past to-buy chair yesterday
'Kipes did not buy a chair yesterday.'/ 'What Kipes did not buy yesterday is a chair.'

*Kipes lo ka-swiim-in kit zoono

A further constraint on embedded sentences in KIHUNG'AN is that the object pronoun -- which may be used as a definitizer or as a pronoun in unembedded sentences -- cannot occur in relative clauses and cannot occur in emphatic constructions.

Schachter and Schachter/Fromkin discuss a tonal change in AKAN which occurs in relative clauses and also in emphatic constructions. In these sentence types certain underlying low tones are replaced by high tones.

(40) a. neutral: ɔ́wɔ̀ Éngìrésì
he-in England
'He is in England.'

b. emphatic: Kòfí nà ɔ́wɔ̀ Engìrésì
Kofi RM he-in England
'It 's Kofi who's in England.'

In SOMALI, on the other hand, a certain agreement pattern links the emphatic constructions to the relative clause. If in SOMALI the subject of a sentence is emphasized, the verb is always in the 3rd person singular, independent of the number of the subject. In other words, agreement does not take place. And this is also true for relative clauses, i.e. in relative clauses the verb is always in the 3rd person singular.

(41) a. neutral: ma?allimíín-tìì wáà̀ yì-màà̀dd-à̀à̀n
the teachers (nom) are come (pl)
'The teachers came.'

b. emphatic: ma?allimíín-tìì báà ti-maadd-a
the teachers (acc) is come (sgl)
'It is the teachers who came.'

All of the above examples clearly show that relative clauses and emphatic constructions are closely related. It has been hypothesized above that the similarities between the two structures are due to the fact that all emphatic constructions are derived from underlying cleft sentences which contain a relative clause. This claim will be investigated in more detail later. Before, however, some data will be presented which supports the claim that emphatic sentences are underlying copula constructions in which an equational relation is established. Even in sentences which superficially look like a non-copulative declarative, traces of an underlying copulative construction are still detectable.

2.6 Signs of an underlying equational structure

Frequently we will find that languages make use of a special marker to indicate emphasis, i.e. in emphatic constructions a morpheme, either bound or free, appears in the immediate neighborhood of the emphasized constituent. In some instances the relation of the emphatic morpheme to other grammatical morphemes, especially to a copula, can be easily established, while in other cases such a relation cannot be readily isolated. The position of the emphatic marker (EM) is generally between the emphasized constituent (C_ε) and the remaining sentence (X). Thus we get the following two combinations:

(a) C_ε EM X
(b) X EM C_ε

Examples are found in many AFRICAN languages, in the MALAYO-POLYNESIAN language group, and in INDIAN languages like BENGALI, HINDI, and MARATHI.

(42) MALAGASY:[11]
a. neutral: tia anao izahay
love you we
'We love you.'

b. emphatic: izahay no tia anao
we EM love you
'It is we who love you.'

(43) TAGALOG:[12]
a. neutral: maganda ang bata
beautiful the child
'The child is beautiful.'

---

[11] Where available both the neutral and the emphatic versions are provided.

[12] Schachter and Otanes (1972) state that the ay inversion forms are not necessarily more emphatic, but rather that (43.b) can be used as a more formal version of (43.a).

b. emphatic: ang bata ay maganda
the child EM beautiful
'It is the child that is beautiful.'

(44) INDONESIAN:
emphatic: menteri lah tuan itu
minister EM man that
'That man is a minister.'

(45) AKAN:
a. neutral: ɔwɔ́ Éngìrésì
he-in England
'He is in England.'

b. emphatic: Kɔ́fí nà ɔ́wɔ́ Engìrésì
Kofi EM he-in England
'It is Kofi who is in England.'

(46) KAREKARE:
a. neutral: tamakən ranho
sheep enter
'The sheep entered.'

b. emphatic: ran na tamakən
enter EM sheep
'It is the sheep that entered.'

(47) NGIZIM:
a. neutral: kadlam papiya
you lied
'You lied.'

b. emphatic: kadlam papiya-n ei
you lied EM you
'It is you who lied.'

(48) BADE:
a. neutral: Dlaagəna ɗaaguraa-gi
Dlaagena called - you
'Dlaagena called you.'

b. emphatic: ɗaaguraa-gi-k Dlaagəna
called-you-EM Dlaagena
'It is Dlaagena who called you.'

While in the above examples the emphatic constituent is either sentence final or sentence initial, there are also languages where the emphatic morpheme emphasizes a preceding constituent that is not sentence final or sentence initial. Two of these languages are LUGBARA, a NILO-SAHARAN language, and MARATHI.

(49) LUGBARA:
emphatic: bă ma á-drîi nï aŋgɔ̀ dɨ̀mà ópi-ɔ̀rɔ̀
they-set-up my brother EM country their chief
'They set up my brother as chief of their country.'

(50) MARATHI:
a. neutral: mi tyana pəyse dein
I them money give
'I give them money.'

b. emphatic: mi tyanats pəyse dein
I them money give
'I give them money.'

Apparently this is also true for BENGALI and HINDI, though all of the examples available to me show subject emphasis and therefore initial position of the emphasized constituent.

(51) HINDI:
a. neutral: ami ɟabo
I go
'I am going.'

b. emphatic: ami-i ɟabo
I EM go
'It is I who am going.'

In some languages it is fairly easy to establish that the emphatic morpheme is either a copula, acts as a copula, is an old form of the copula, or is a contraction of the copula and the relative clause marker. In HAUSA, for instance, the emphatic nē particle also occurs in constructions which indicate equivalence:

(52) Audu nē malami
Audu is teacher
'Audu is a teacher.'

In INDONESIAN, the lah particle is sometimes used as a copulative element, and in LUGBARA nï occurs after nominal and adjectival predicates without emphatic function.

(53) LUGBARA:
ágú ʔdàri ópí (-mvá)-nï
man this chief's son is
'This man is a chief('s son).'

A language where the emphatic morpheme appears to be the result of a merger between the copula and the relative pronoun is AKAN. In AKAN the so-called emphatic particle is na which emphasizes a preceeding

constituent. According to Balmer, Grant (1929) na is a merger of ne, an older form of the copula, and of the relative clause introducer a. While in the above-discussed languages it can be shown that the emphatic marker is identical with the copula, such a relationship is not as easily established in other languages. One of the reasons is that most of the languages that use an emphatic marker in a cleft construction do not have a copula in other equational sentences. In some languages, for instance, in some of the CHADIC and CUSHITIC languages, the constraints on the verb form of the emphatic sentence reveal that the emphatic sentence has an underlying cleft construction even though the nature of the emphatic marker cannot be clearly established. Further research will show whether or not all emphatic morphemes can be related in some form or other to a copula. If not, this will not invalidate the hypothesis argued for in this paper. All it will say is that some languages mark the focus constituent by using a special morpheme, comparable to the way some languages mark the focus constituent by using phonological means like increased loudness or a change in tone.

Another sign of an underlying equational structure is case assignment, as for instance in SOMALI. In SOMALI the emphasized constituent is always in the accusative which according to Tucker and Bryan is the unmarked case of SOMALI.

(54) a. Subject emphasis
nin-kíi baà lị̀bá̰à̰h-íí arka
man-acc. is lion-acc. sees
'It is the man who sees the lion.'

b. Object emphasis
nín-kíí baà lị̀bá̰à̰h-íḭ́ àrkàà
man-acc. is lion-nom. sees
'It is the man that the lion sees.'

This fact can be explained if we assume that emphatic sentences are derived from underlying equational sentences because in SOMALI the predicate nominal of an equational sentence is always in the accusative case. On the other hand, if we do not make the above assumption, then we need an ad hoc constraint which insures that the emphasized constituent is always marked for accusative.

Another sign of an underlying equational construction is the presence of a pause between the known information and the focused constituent. In the examples below a pause is indicated by a comma.

(55) TSWANA:

a. neutral: rrê orêkilê dikgômo maabane
my-father he-bought cattle yesterday
'My father bought cattle yesterday.'

b. emphatic: orêkilê dikgômo maabane, rrê
he-bought cattle yesterday, my-father
'It is my father who bought the cattle yesterday.'

(56) TELUGU

a. neutral: raamaaraawu jarmaniininci kotta kaaru teppincEEDu
Rama Rao from-Germany new car got
'Rama Rao got a new car from Germany.'

b. emphatic: raamaaraawu jarmaniininci teppincindi, kotta kaaru
Rama Rao from-Germany got new car
'It is a new car that Rama Rao got from Germany.'

It is fairly common to use a pause instead of a copula element to indicate an equational relation, e.g. in RUSSIAN, VIETNAMESE and TARTAR. The pause demarcates the line between the subject or known information and the predicate which contains the new information. This pause becomes especially important in longer sentences where the demarcation line is not immediately obvious from the structure of the sentence.

The claim that (55.b) and (56.b) are equational structures is further supported by the position of the emphasized constituent in sentence final position. The position of the emphasized constituent in a cleft sentence is discussed in more detail in section 4.

3. The underlying structure of a cleft sentence

Earlier in this paper it was hypothesized that all emphatic constructions are derived from underlying cleft sentences and evidence from many languages was presented to show that the underlying cleft structure is still traceable in the surface structure, even though cleft constructions can undergo considerable reductions in certain languages. In the following I will discuss the structure of the underlying cleft sentences and the rule component required to generate the various surface manifestations of emphatic constructions.

In recent linguistic writings cleft sentences have been fairly widely investigated.[13] Jesperson (1949) suggested for instance that cleft sentences are derived from underlying simples sentences by the insertion of it is → WH:

[13] See Jespersen, 1949, 1969; Lees, 1963; Chomsky, 1970; Bach/Peters 1968; Akmajian, 1970; Schachter, 1973. For a discussion of Bach/Peters, Chomsky, and Akmajian, see Harries, 1972. For a discussion of Schachter's proposal, see below.

(57) The man called. → It is the man who called.

As Lees (1963) points out, such a derivation of a complex structure from a simplex one does not account for the fact that we can have different tense and aspect markings in the two verbs of the cleft and also that we can get negation in one or the other or in both of the sentences.

(58) It must have been my mother who flew to New York.
(59) It is not my mother who flew to New York.
(60) It is my mother who didn't fly to New York.
(61) It is not my mother who didn't fly to New York.

On the basis of the above data it has therefore been generally assumed that cleft sentences are derived from underlying complex structures. Less agreement is found, however, when we look at the kinds of underlying structures proposed. Lees suggested that clefts are derived from it + AUX + be + complement structures. Chomsky and Akmajian assumed a structure in which the head noun of the subject relative clause is the it morpheme which shows up in sentences like (58) to (61),[14] while Bach/Peters assume a head noun the thing, thus excluding cleft constructions with the one. That the it morpheme plays such a prominent role in nearly all of the above analyses seems to be mainly due to the fact that all of these linguists based their analyses on data from ENGLISH. In ENGLISH cleft constructions with it are very common, but they do not exist in most other languages. That is, even though as far as I could determine all languages do have cleft constructions of the form a) the one who ... is X, only relatively few languages have a corresponding form like b) it is X who ... . And whenever a language has b), it also has a).

The conditions under which a neutral pronoun has to be inserted will be discussed in more detail in the following section. It will be argued that certain languages require the insertion of a pro-subject whenever the underlying subject nounphrase was extraposed. In other words it is argued that the it-pronoun is predictable and consequently is not part of the underlying structure. Besides the fact that an analysis that places the it-pronoun into the deep structure does not account for the predictability of this pronoun, it also fails to account for the paraphrase relationship between sentences like (62) and (63) below.

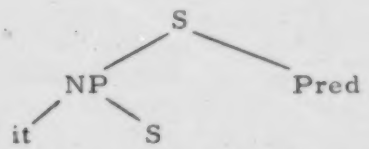
(62) John was the one who helped us.
(63) It was John who helped us.

---

[14] The major difference between Akmajian and Chomsky is that the former assumes that the underlying predicate dominates the focus constituent, while the latter assumes that the predicate dominates a dummy symbol which is replaced by one of the noun phrases of the subject relative clause with the help of an extraction rule.

(62) and (63) are paraphrases because they contain identical presuppositions and make an identical assertion and consequently should be derived from the same underlying form. However, no analysis which assumes a structure as in (64) will be able to relate (62) and (63) in a non-ad hoc way.

(64)



Based on the above evidence I will therefore assume that in all languages the underlying structure of cleft sentences contains a head noun of the form the one, the person, the he, the thing[15] (the place, the time -- not discussed here). I am not sure about the exact feature specification of the neutral head noun. It will always be third person; furthermore, in some languages it will be marked for number (ENGLISH, GERMAN, FRENCH), humanness (ENGLISH, MANDARIN CHINESE), gender (GERMAN, FRENCH) and possibly a few other features. However, given that it is a neutral noun, it will not contain any unique semantic features, that is features that would differentiate it from every other noun in the language.

Problematic for setting up a universal underlying structure is the linear order in which the constituents occur. In other words, the question is whether there is one universal linear order, whether there are different underlying orders which in turn are reflected in the different surface orders or whether there is no underlying syntactic order at all and surface orders are imposed by a set of language specific ordering rules. I tend to favor the latter position. However, this paper is not the place to argue these questions. All I will therefore say at this point is that the underlying structure of all cleft constructions contains a subject and a predicate. The subject is composed of a neutral head and a relative clause which restricts the large set specified by the head noun. The subject contains the known information, while the predicate contains the focus constituent which is the new information.[16] Whether or not

---

[15] In VIETNAMESE the neutral head noun is identical with the classifier of the focus constituent.

ông Ba ông ấy hay đi Dalat lắm.
Mr. Ba nr.cl. that often go Dalat very
'Mr. Ba is the one who very often goes to Dalat.'

[16] 'New information' here is meant in the sense used by Halliday (1967): "The focus of the message, it is suggested, is that which is presented by the speaker as being new, textually (and situationally) non-derivable information ... But the non-predictability of the new does not necessarily imply factually new information; the newness may lie in the speech function, or it may be a matter of contrast with what has been said before or what might be expected." (pp. 205-6).

a copula is part of the underlying form is an open question. Bach (1967) argued that the copula should be inserted by a transformation for the following reasons:

a) The copula is predicatable.
b) It complicates the grammar if the copula is in the deep structure.
c) There are languages that do not have a copula.

a) seems to hold for ENGLISH but will have to be tested for other languages. b) is difficult to verify as it subsumes the existence of a complete grammar for a given language and of an evaluation measure. c) is true for a few languages like TAGALOG and BLACKFOOT. More common is it to find languages in which the copula can be absent under certain condition. Most commonly it can be absent in the present tense but not in the other tenses (e.g. DRAVIDIAN languages, RUSSIAN, HEBREW, JAPANESE, BANTU languages). In BENGALI it can be absent in the present tense if the sentence is affirmative.

The opposite stand from that of Bach was taken by Ross (1969a, 1969b). In the first paper Ross claims that 'be' is a true verb which is present in the deep structure and is marked for the feature $[\ ^{+V}_{-Adj}\ ]$.[17] No evidence is presented in support of this claim. In the second paper Ross argues that auxiliaries are underlying main verbs and he assumes that have and be belong into this class. There is, however, evidence from negation that even though modals like can and must appear to be underlying main verbs, this claim cannot be extended to auxiliaries like have and be.[18] The claim that be is a main verb is also weakened if we look at other languages besides ENGLISH. Outside of the INDO-EUROPEAN family we rarely find copulas that behave like true verbs. That is, very often the copula is an invariant morpheme which does not undergo the inflectional processes that true verbs undergo. And frequently the copula will have a unique negative form, i.e. negation of the copula and of true verbs does not take the same form. Furthermore, the copula is not always found in the position of the verb. Ferguson (1972) notes that in languages that develop a pro-copula,[19] this form will tend to occur between the subject and the complement, even in SOV languages like BENGALI or in VSO languages like the SEMITIC ones.

(65) BENGALI:
a. cheliṭi chatro
boy student
'The boy is a student.'

---

[17] For a discussion see Darden, 1969.

[18] For a discussion see Harries, 1973a.

[19] Pro-copula is defined as "an expression whose primary function is non-copulative, used in place of zero in a copulative sentence." (Ferguson, 1972: 95).

b. cheleṭi {hɔ̌e / holo / hocce} chatro
boy is student
'The boy is a student.'

c. cheleṭi chatro nɔ̌e
boy student not-is
'The boy is not a student.'

In (65.b) the pro-copula occurs in medial position while in (65.c) the negative copula occurs in the final verb position.

(66) ARABIC (CLASSICAL)
a. Aliyun mu?alimun
Ali a-teacher
'Ali is a teacher.'

b. Aliyun huwa mu?alimun
Ali he a-teacher
'Ali is {a / the} teacher.'

c. Aliyun huwa lā mu?alimun
Ali he not a-teacher
'Ali is not a teacher.'

SEMITIC languages like ARABIC can optionally use a third person pronoun or a demonstrative as pro-copula which will occur in medial position even though ARABIC is a VSO language, and the negative particle lā follows the pro-copula.

All of the above outlined facts indicate that within a universal framework the copula cannot be considered a true verb as Ross proposes. On the other hand, if the copula is inserted by rule, there is no node to attach it to. I therefore assume that a pro-node is present in the underlying form of equational sentences. This node is marked for being stative and it is directly dominated by S. The vagueness of the node allows for the great variation in its surface realization across languages. Some languages will not realize it at all; others will realize it before nouns but not before adjectives, or in the past tense but not in the present tense. In ENGLISH it behaves in many respects like a verb, in other languages it is a grammatically unique element that does not share any feature with the verb or any other constituent type.

4. The surface order of cleft sentences

Above I said that no underlying linear order will be assumed for cleft sentences. However, cleft sentences do have a surface order and I want to argue that this order depends on the language type. That

is, we can make the following prediction where S = subject NP, Cop = copula, and F = focus NP.[20]

	Language type	Basic cleft construction
a.	SVO	S Cop F
b.	SOV	S F Cop
c.	VSO	Cop S F

While examples of a. and b. are readily available, this is not true for c. I was not able to find data of a VSO language with a copula in initial position. As mentioned earlier, ARABIC uses a third person pronoun as pro-copula in medial position. MAASAI, another VSO language, has zero copula and so does TAGALOG. Unfortunately, I was not able to get any data for LOTUHO, an EAST AFRICAN language with VSO order and an overt copula. It is essential, however, that in all three language types the position of the focus nominal corresponds to the position of the object in a non-cleft sentence, i.e. in SVO and VSO languages the focus nominal occurs in final position while in SOV languages it occurs in medial position. The underlined constituents are the object and the focus constituent respectively.

(67) SVO -- S Cop F
a. John bought a book.
b. What John bought is a book.

(68) SOV -- S F Cop
JAPANESE:
a. John wa Mary o butta
John subj. m. Mary obj.m. hit
'John hit Mary.'

b. Mary o butta no wa John da.
Mary obj.m. hit RM subj.m. John is
'The one who hit Mary is John.'

(69) VSO -- (Cop) S F
ARABIC (COLLOQUIAL):
a. kataba Ali ar-risāla
wrote Ali the-letter
'Ali wrote the letter.'

b. elli kataba ar-risāla huwa Ali
RM wrote the-letter he Ali
'The one who wrote the letter is Ali.'

---

[20] VOS languages like MALAGASY were not included in this discussion for lack of information about this language type. It appears, however, that VOS languages should have a basic Cop F S cleft order.

Above it has been argued that if we know the position of the object in a neutral sentence, we can predict the basic position of the focus constituent in a cleft sentence. Later on it will be shown that the position of the object is also the basic position of an emphatic constituent in a non-equational sentence,[21] even if the emphatic constituent is the surface subject of that sentence. However, before investigating non-neutral emphatic word orders in surface simplexes, I will discuss the rule component required to transform underlying cleft constructions into emphatic surface forms, which may or may not be clefts.

5. The rule component

The rules discussed below are assumed to be part of the universal grammar though not all of the languages will have all of the rules. Constraints on the individual rules are assumed to be language specific. The five major rules involved are: 1) deletion of the neutral head noun, 2) deletion of the relative clause introducer, 3) deletion of the copula element,[22] 4) extraposition, and 5) predicate lowering.

5.1 Deletion rules

5.1.1 Deletion of neutral head noun. Given that the head noun in a cleft construction does not carry any information which is not also present elsewhere in the sentence,[23] it is not surprising that in most languages the head noun is optionally deletable at least in some environments (see examples below). Whether there are languages in which the head noun is obligatorily deleted in all environments I was not able to establish. In some languages like ENGLISH and FRENCH, the head noun is obligatorily deleted if subject extraposition applies, while in languages like HUNGARIAN and KIHUNG'AN head deletion is optional even after subject extraposition.

(70) a. ? It's my friend, the one who came.
b. It's my friend who came.

---

[21] I shall refer to these non-equational sentences as simplex sentences even though they could obviously also contain embedded clauses.

[22] The rule that deletes the [static] node in certain languages under certain conditions will not be discussed. Some of the conditions under which the copula node is deleted were mentioned above.

[23] That is, all of the information on number, gender, humanness, etc. which can be present in the head noun can be predicted from the focus constituent, as there exist cooccurrence restriction between the head noun and the focus constituent to the effect that all of the features of the head noun must also be found in the focus constituent. For instance, a head noun marked for male can only be placed into an equivalence relation with another male noun.

(71) FRENCH:

a. ?C'est mon ami celui qui est venu.
? 'It's my friend the one who came.'

b. C'est mon ami qui est venu.
'It's my friend who came.'

(72) HUNGARIAN:

a. Nem John volt az aki New York-ba repült.
Not John was that who New York-to flew
'It was not John who flew to New York.'

b. Nem John volt aki New York-ba repült.
Not John was who New York-to flew
'John was not the one who flew to New York.'

(73) KIHUNG'AN:

a. (kwe) kít kiim ki a-swiim-in Kipes zoon
(is) chair thing that bought Kipes yesterday
'The chair is the thing that Kipes bought yesterday.'

b. (kwe) kit ki a-swiim-in Kipes zoon
(is) chair that bought Kipes yesterday
'it is the chair that Kipes bought yesterday.

While some languages require the obligatory deletion of the neutral head noun in some environments, some languages do not permit the deletion of the head noun under certain conditions. In MANDARIN CHINESE, for instance, head deletion is constrained in long sentences. The presence of the head helps to delineate the subject and to establish the equational relation. In ENGLISH the deletion of the head noun is not possible if it is human, e.g.:

(74) a. The one who sang was Caruso.
b. *Who sang was Caruso.

However, as will be seen later, a structure like (74.b) above is required to generate surface simplexes from underlying cleft sentences (see predicate lowering). It therefore appears that even in ENGLISH head deletion can take place in any environment but that either subject extraposition or predicate lowering have to apply if the head noun has been deleted.

5.1.2 Deletion of the relative clause marker. The deletability of a relative clause marker does not seem to depend on the nature of the marker itself. That is, there is no evidence that the marker is more often deletable if it is invariant than if it is a pronoun. For instance, in ENGLISH the relative pronoun can be deleted if certain conditions are met (e.g. if the relativized constituent is not the subject of the relative clause).

(75) a. The one (who) I met is John's brother.
b. The one (who is) studying in the library is a friend of mine.

Deletion of the relative pronoun is also possible in BANTU languages like KIHUNG'AN:

(76) a. (kwe) kit ki a-swiim-in Kipes zoono
(is) chair that bought Kipes yesterday
'It is a chair that Kipes bought yesterday.'

b. (kwe) kit Kipes a-swiim-in zoono
(is) chair Kipes bought yesterday
'It's a chair that Kipes bought yesterday.'

(77) a. (kwe) Kipes wu a-swiim-in kit zoono
(is) Kipes who bought chair yesterday
'It's Kipes who bought a chair yesterday.'

b. (kwe) Kipes a-swiim-in kit zoono
(is) Kipes bought chair yesterday
'It's Kipes who bought a chair yesterday.'

In GERMAN, on the other hand, relative pronouns are not deletable.

(78) a. Derjenige, den du trafst, ist Peters Bruder.
'The one whom you met is Peter's brother.'

b. *Derjenige du trafst ist Peters Bruder.

Among the languages with an invariant relative clause marker, MANDARIN CHINESE does not allow the deletion of the marker de. In MALAGASY, on the other hand, the relative marker izay does not show up in emphatic constructions. Both head and relative marker deletion can apply in some languages.

(79) TAGALOG:
a. si Rosa ang siyāng nakita niya
Constr.m. Rosa nom. he-RM saw he
'Rosa is the one who he saw.'

b. si Rosa ang nakita niya
Constr.m. Rosa nom. saw he
'Rosa is the one who he saw.'

(80) KIHUNG'AN:
a. (kwe) kit kiim ki a-swiim-in Kipes zoono
(is) chair thing that bought Kipes yesterday
'It's a chair that Kipes bought yesterday.'

b. (kwe) kit Kipes a-swiim-in zoono
(is) chair Kipes bought yesterday
'It's a chair that Kipes bought yesterday.'

5.2 Reordering rules

5.2.1 Extraposition or fronting. Up to now we have discussed those rules that delete certain elements from an underlying cleft construction. Another type of change that cleft sentences can undergo is that constituents can be moved around and reordered resulting in sentences like (82) and (83).

(81) The one who stole the cheese is Frank.
(82) Frank is the one who stole the cheese.
(83) It is Frank who stole the cheese.

Both (82) and (83) are derived from (81) by movement rules. The effect in both (82) and (83) is that the focus constituent is brought closer to the front of the sentence. It is a well-known fact that frontal position is a position of emphasis in most languages. The question is: how does the focus constituent get into that position and how do we account for the difference between (82) and (83)? Logically, there are two possible alternatives that will convert (81) into (82) and (83), i.e. either the focus constituent is fronted, or the known information is placed at the end of the sentence. The first rule I will call focus fronting, the second extraposition. It is possible that either both of these rules exist in natural language or that only one or the other one does. If we assume the existence of both rules, we can account for the difference between (82) and (83) above.

(84)	a.	Basic order:	The one who stole the cheese is Frank.
	b.	Fronting:	Frank the one who stole the cheese is.
	c.	Copula placement:	Frank is the one who stole the cheese.
(85)	a.	Basic order:	The one who stole the cheese is Frank.
	b.	Extraposition:	is Frank (the one) who stole the cheese.
	c.	Pro-subj. insertion:	It is Frank who stole the cheese.

If, on the other hand, we assume that focus fronting is the only reordering rule that applies to cleft sentences, then the derivation of (83) above requires one additional rule, namely copula fronting.

(86)	a.	Basic order:	The one who stole the cheese is Frank.
	b.	Focus fronting:	*Frank (the one) who stole the cheese is.
	c.	Copula fronting:	is Frank (the one) who stole the cheese.
	d.	Pro-subj. insertion:	It is Frank who stole the cheese.

On the other hand, one might want to say that focus fronting can optionally move both the focus constituent and the copula simultaneously, converting (86.a) directly to (86.c). The existence of such an optional constraint on fronting -- though technically feasable -- appears nevertheless ad hoc and is not corroborated by any additional evidence. But even if we assumed such a constraint, the copula placement rule in derivation (84)

would still be required, though not the copula fronting rule assumed in derivation (86). If, on the other hand, we assume that extraposition is the only movement rule that applies to cleft sentences, no copula placement or fronting rule is required.

(87) a. Basic order: The one who stole the cheese is Frank.
b. Extrap. (subj. NP): is Frank (the one) who stole the cheese.
c. Pro-subj. insertion: It is Frank who stole the cheese.

(88) a. Basic order: The one who stole the cheese is Frank.
b. Extrap. (copula): The one who stole the cheese Frank is.
c. Extrap. (subj. NP): Frank is the one who stole the cheese.

The difference between derivation (87) and (88) is that in (87) extraposition first applied to the subject NP, while in (88) it first applied to the copula. This means that extraposition can apply either to the constituent that is sentence initial as in (87), or to the constituent that precedes the focus constituent (pre-focus constituent) as in (88), the common condition being that the extraposed elements contain known information. The assumption that extraposition can apply at two different points of the string will appear ad hoc at this point of the analysis. It will, however, be shown later during the discussion of marked word orders in emphatic simplex sentences that we can only account for the many and seemingly unsystematic variations found in emphatic word order if we make the above assumptions about extraposition. That is, we have to assume that: 1) extraposition is the only reordering rule that changes the linear order of constituents in an emphatic sentence; 2) extraposition applies to old or presuppositional material only; 3) extraposition can apply to the constituent which is in sentence initial position, or to the constituent that precedes the focus constituent.

5.2.2 Extraposition. For the moment I will not present any further support for the above assumptions but will rather show how extraposition accounts for cleft sentences in various languages. In some of the above derivations a rule called pro-subject insertion placed the neutral pronoun it into the position of the subject NP after the latter had been extraposed. This rule is not restricted to ENGLISH. It also applies in RUSSIAN (éto 'this'), FINNISH (se 'this'), FRENCH (ce 'this'), HEBREW (ze 'this') and GERMAN (es/das 'it/that').

(89) HEBREW:
a. Basic order:
ze she pagash et axiv haya Yonatan
this RM met DOM brother-his was John
'The one who met his brother was John.'

b. Extraposition (subject NP):
haya Yonatan (ze) she pagash et axiv
was John (this) who met DOM brother-his

c. Pro-subject insertion:
ze haya Yonatan she pagash et axiv
this was John who met DOM brother-his
'It was John who met his brother.'

(90) FRENCH:

a. Basic order: Celui qui est venu est François.
'The one who arrived is François.

b. Extraposition: est François (celui) qui est venu
is François (one) who arrived

c. Pro-subj. insertion: C'est François qui est venu.
'It is François who arrived.

As far as I could establish, pro-subject insertion is restricted to SVO languages where it takes place if the subject NP is extraposed.

(91) a. Basic order: S - Cop - F
b. Extraposition: Cop - F - S
c. Pro-subj. insertion: PS - Cop - F - S

A probable explanation for the pro-subject insertion rule is that a surface order as in (91.b), where the verbal element is sentence initial, signals a yes-no question in many languages. In order to avoid ambiguity, a pro-subject is inserted. This hypothesis is supported by the fact that no pro-subject insertion takes place in SVO languages in which initial position of the verb does not signal a question.[24] This is for instance true for BANTU languages like SWAHILI and KIKUYU, and for MANDARIN CHINESE. In all three of these languages, (91.b) above is a possible surface construction.

(92) SWAHILI:

a. mtu aliye fika ni mwalimu
man who arrived is teacher
'The one who arrived is a teacher.'

b. ni mwalimu mtu aliye fika
is teacher man who arrived
'It is a teacher who arrived.'

---

[24] The hypothesis that only those languages have pro-subject insertion in which initial position of a verb signals a question is also supported by the fact that VSO languages have no pro-subject insertion. Obviously, verb initial position cannot signal a question in VSO languages.

(93) KIKUYU:

a. nĩn gahaica kĩrĩma kĩu nĩ rũciũ
I shall-climb that hill is tomorrow
'(The time) when I shall climb that hill is tomorrow.'

b. nĩ rũciũ ngahaica kĩrĩma kĩu
is tomorrow I-shall-climb that hill
'It is tomorrow that I shall climb that hill.'

(94) MANDARIN CHINESE:

a. kàn-jiàn nà-ge nán-rén de (rén) shìh Yohàn
saw that-cl. male-person RM (person) is John
'The one who saw that man is John.'

b. shìh Yohàn kàn-jiàn nà-ge nán-rén
is John saw that-cl. male-person
'It is John who saw that man.'

That SOV languages do not have pro-subject insertion is due to the fact that subject extraposition will not result in a sentence initial position of the copula.

(95) a. Basic order: S F Cop
b. Extraposition: F Cop S

Examples of (95) are:

(96) HUNGARIAN:

a. Basic order: Az aki New York-ba repült János volt.
that who New York-to flew John was
'The one who flew to New York was John.'

b. Extraposition: János volt az aki New York-ba repült.
John was that who New York-to flew
'John was the one who flew to New York.'

(97) AMHARIC:

a. Basic order: ya maṭṭaw saw əssu naw
RM came man he is
'The one who came is he.'

b. Extraposition: əssu naw ya maṭṭaw
he is RM came
'He is the one who came.'

It was argued earlier that extraposition can apply to the sentence initial constituent or to the pre-focus constituent, provided of course that these constituents represent old information. Thus, given an SVO language, we would expect that the following variations occur:

(98) a. Basic order: S Cop F
b. Extraposition sentence initial: Cop F S
c. Extraposition: ?F S Cop

(99) a. Basic order: S Cop F
b. Extraposition pre-focus: ?S F Cop
c. Extraposition: F Cop S

The question marks indicate that I did not find any languages that have either (98.c) or (99.b) as a variation of the basic order. I can only speculate why this should be so. (98.c) and (99.b) are identical in that in both orders the copula occurs in final position, an order which S Cop F (or SVO) languages seem to avoid. However, it is conceivable that a "free" word order language would also have (98.c) and/or (99.b) as alternate orders of the basic one. While a basic S Cop F order allows for four variations, S F Cop orders (SOV languages) are much more restricted, given that the focus constituent is in second position.

(100) a. Basic order: S F Cop
b. Extraposition: F Cop S

(100.b) is the only variation found in SOV languages (see the HUNGARIAN and the AMHARIC examples above).

VSO languages, on the other hand, should allow as wide a range as SVO languages:

(101) a. Basic order: Cop S F
b. Extraposition sentence initial: S F Cop
c. Extraposition: F Cop S

(102) a. Basic order: Cop S F
b. Extraposition pre-focus: Cop F S
c. Extraposition: F S Cop

Unfortunately, I have very limited data from VSO languages. In ARABIC, as mentioned earlier, a. does not occur but rather S Cop F, if we want to assume that the third person pronoun functions as a copula. As a variation, c. is found in ARABIC. But further research is definitely needed before any substantive claims can be made.

As pointed out earlier, not all of the forms occur in all languages. Furthermore, there are languages that apparently do not have the extraposition rule, at least not for nominals. Thus in JAPANESE, an SOV language, the expected alternative with the focus constituent in initial position is not grammatical, nor is any other word order variation.

(103) a. Mary o butta hito wa Bill da
Mary obj.mk. hit person subj.mk. Bill is
'The one who hit Mary is Bill.'

b. *Bill da Mary o butta hito wa
Bill is Mary obj.mk. hit person subj.mk.

In MANDARIN CHINESE, extraposition can only apply to the sentence initial and not to the pre-focus constituent and furthermore, the focus constituent has to correspond to the subject of the relative clause. Thus both (104.b) and (105.b) are ungrammatical.

(104) a. kàn-jiè nà - ge nán - rén de shìh Yohàn
saw that- cl. male-person RM is John
'John is the one who saw that man.'

b. *Yohàn shìh kàn-jiè nà-ge nán - rén de
John is saw that-cl. male-person RM

(105) a. dzúo-tien wǒ mǎi de shìh shu
yesterday I bought RM is book
'What I bought yesterday is a book.'

b. *shìh shu dzúo-tien wǒ mǎi de
is book yesterday I bought RM

In HEBREW, too, extraposition can only apply to the sentence initial constituent, as an F Cop S order is ungrammatical.

(106) a. zeh sheshuh pagash etmol haya axiv
this RM met yesterday was brother-his
'The one he met yesterday was his brother.'

b. *axiv haya zeh sheshuh pagash etmol
brother-his was this RM met yesterday

c. ze haya axiv sheshuh pagash etmol
{this / it} was brother-his RM met yesterday
'It was his brother whom he met yesterday.'

The above data suggests that if a language constrains the extraposition rule, it will exclude first the pre-focus application. In other words, sentence initial application is the more basic one. I did not find any language in which extraposition can apply to the pre-focus constituent and not to the sentence initial one. In some languages it is difficult, however, to decide which kind of extraposition took place. This is true for all those languages that have a zero copula like in the following example from DIOLA-FOGNY.[25]

[25] I expect, however, that additional data from DIOLA-FOGNY would allow us to decide one way or the other.

(107) a. nijuk -ɛ ebe
I-saw-ɛ cow
'What I saw is a cow.'

b. ebe nijuk-ɛ
cow I-saw-ɛ
'A cow is what I saw.'

5.3 Predicate lowering

All of the surface manifestations of emphatic constructions discussed so far involve complex equational sentences which are composed of a minimum of two S-nodes, the matrix S and the relative clause S. It is, however, a fact that some languages have emphatic constructions that are clearly non-equational and non-complex, e.g.:

(108) John saw { Frank / him / *he } yesterday (but not Bill).

(109) GERMAN:
Ich habe nicht { den Tisch / *der Tisch } gekauft (sondern den Stuhl).

'I didn't buy the table, (but rather the chair).'

Support for the claim that John saw Frank and Ich habe nicht den Tisch gekauft are simplexes comes from case assignment. Both in ENGLISH and in GERMAN the emphasized constituent is marked for accusative case in the reduced sentence. In a cleft sentence, on the other hand, the predicate noun phrase is marked for nominative.

(110) { He / Frank / *He } is the one John saw.

{ Der Tisch / *Den Tisch } is nicht das, was ich gekauft habe.

'The table is not that, what I bought.'

Consequently, if the emphasized constituents in (108) and (109) were the predicates of a higher S, they should display the case marking that would be assigned to them in that S, i.e. nominative, but they don't.[26] Another sign that a complex cleft sentence has been reduced to a simplex is the order in which the constituents occur. That is, in a cleft construction the focus constituent is either to the left or to the right of the known information, but never surrounded by it. In other words,

---

[26] SOMALI, on the other hand, is a language in which the case marking of the emphasized constituent is retained in the simplex. For a discussion see below.

there are two possible orders (not considering the copula), one in which the focus constituent follows the known information, e.g. (112), and one in which it precedes the known information, e.g. (113), but not (114) in which the focus constituent is surrounded by old information. (114) cannot be a cleft, i.e. an equational construction.

(112) The one John watched yesterday is my friend.
(113) My friend is the one John watched yesterday.
(114) John watched my friend yesterday.

However, even though (114) is not a cleft, I want to argue that it was derived from a cleft like (112). What is the evidence for this claim? First, there is the semantic paraphrase relation.

(115) a. What I bought is the book.
b. I bought the book.

(115.a) and (115.b) are semantically equivalent, i.e. both sentences presuppose that I bought something and both assert that the thing bought is a book. Besides the semantic evidence there is, however, also considerable syntactic evidence which supports that (115.b) is derived from (115.a). This evidence comes from negation and question formation. For instance, in an emphatic construction in languages like HUNGARIAN, GERMAN, NAVAHO, and RUSSIAN, the negative particle may occur before the emphasized constituent instead of in the regular sentence negation position before or after the verb.[27]

(116) GERMAN:
a. neutral: Mein Vater ist nicht gekommen.
'My father didn't come.'

b. subj.emph.: Nicht mein Vater ist gekommen, (sondern mein Bruder).
'My father didn't come, (but my brother (did)).'

(117) RUSSIAN:
a. neutral: Ja ne edu v London.
I not going to London
'I am not going to London.'

b. subj.emph.: Ne ja edu v London.
Not I going to London
'It is not I who is going to London.'

---

[27] In languages like ENGLISH and MANDARIN CHINESE, on the other hand, reduction to a simplex is not possible if the highest S of the cleft sentence is negated. That is, both ENGLISH and MANDARIN CHINESE cannot reduce a sentence like (a) to a sentence like (b):

(a) The one who bought the oranges is not Bill.
(b) *Not Bill bought the oranges.

(118) HUNGARIAN

a. neutral: János nem vett könyvet.
John not bought book
'John didn't buy a book.'

b. subj.emph.: Nem János hanem Mary vett egy könyvet.
Not John but Mary bought a book
'It is not John but Mary who bought a book.'

(119) NAVAHO:

a. neutral: John adadā ado niyada
John yesterday not arrive-neg.
'John didn't arrive yesterday.'

b. emphatic: do adadā John niyada, hadashi
not yesterday John arrive-neg, but some other time
'John didn't arrive yesterday, but (he arrived) some other time.'

It has been claimed, at least for GERMAN[28] that the position of the negative particles in sentences like (116.b) is due to 'constituent' negation. Presumably, this means that not the sentence but a single constituent is negated. However, in that case we have to ask why constituent negation of this kind can only occur before a constrastively-emphasized constituent. That is, a sentence like (116.b) is ungrammatical if the speaker does not intend to contrastively emphasize the subject. On the other hand, true word or constituent negation as in unmöglich 'impossible', mistrauen 'mistrust', Nichtraucher 'non-smoker', etc. can occur without any such constraint.

(120) Der Nichtraucher ist gekommen.
'The non-smoker arrived.'

(121) *Der Nichtraucher ist gekommon, sondern der Abstinenzler.
'The non-smoker arrived, but rather the teetotaller.

In (120) Nichtraucher is clearly word or constituent negation. It cannot occur in an emphatic environment, e.g. (121). Furthermore, the position of nicht, 'not/non' in (120) differs from the position of nicht in (116.b) in that in (120) nicht is part of the noun and therefore follows the article, while in (116.b) it is outside of the noun and has to precede the article.

(122) *Mein nicht Vater ist gekommen.
'My non/not father arrived.'

Consequently, given that the negative particle in sentences like (116.b) cannot be constituent negation, we have to assume that it is sentence

---

[28] See Ibañez, 1970.

negation. How, then, do we account for the position of the negative particle in these sentences? We might claim that languages that allow this particle to occur before the emphasized constituent have a rule which simply says: place the negative particle before the emphasized constituent. Such a rule would convert a sentence like (116.a) into a sentence like (116.b) if the subject is somehow marked for emphasis. However, such an analysis will not account for the fact that contrastively-stressed sentences can contain two negative particles, e.g.:

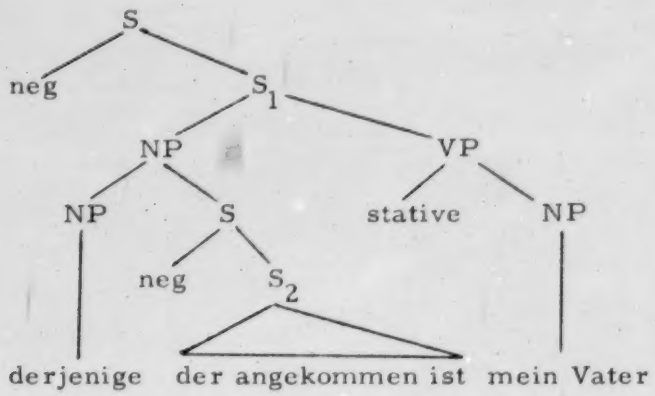
(123) Nicht mein Vater ist nicht angekommen, (sondern mein Bruder ist nicht angekommen).

Not my father didn't arrive, (but my brother didn't arrive).
'It is not my father who didn't arrive, (but it is my brother who didn't arrive.'

(124) Nem John nem repült New York-ba, hanem Bill nem repült New York-ba.
Not John not flew New York-to, but Bill not flew New York-to.
'It is not John who didn't fly to New York, but it is Bill who didn't fly to New York.'

If both occurrences of nicht in the surface simplex Nicht mein Vater ist angekommen constitute sentence negation, then we have to assume that this sentence is derived from an underlying complex structure as in (125) below:

(125)



In a cleft sentence, negation of S_1 always implies that the equivalence relation between the subject NP and the predicate NP is rejected. This implication does not exist if S_2 is negated. (125) undergoes negative lowering which will result in a surface form like (126) below.

(126) Derjenige, der nicht angekommen ist, ist nicht mein Vater, (sondern...).
'The one who didn't arrive ist not my father, (but ...).'

(126) can be further reduced to (123). Thus the assumption that (125) underlies a sentence like (123) allows us to account -- without any new rules -- for the occurrence of two sentence negations in (123) and for the position of the negative particles before the emphasized constituent. As (126) shows, sentence negation in an equational structure is always placed before the predicate constituent, e.g.

(127) Hans ist kein Lehrer.
Hans ist not-a teacher

(128) Marika ist nicht dumm.
Marika ist not stupid.

Above we have seen that evidence from negation indicates that surface simplexes with an emphatic constituent are derived from underlying complex sentences. Further evidence of the same kind is provided by some languages in which the question particle can be associated with the emphatic constituent. In FINNISH, for instance, the question marker ko / kö is generally attached to the verb, e.g. (129). In an emphatic question, speakers of FINNISH have two choices. They can either retain the structure of the neutral question and stress the emphasized constituent, or they can place the question particle on the emphasized constituent.

(129) a. neutral: Näkikö äiti Billin?
saw-Q mother Bill-acc.
'Did mother see Bill?'

b. subj.emph.: Näkikö äiti Billin?
saw-Q mother Bill-acc.
'Did mother see Bill?'

c. subj.emph.: Äitikö Billin näki?
mother-Q Bill-acc. saw?
'Did mother see Bill?'

The same pattern is also found in TURKISH, where the question marker mi, which generally precedes the verb, can optionally precede the emphatic constituent.

(130) a. neutral: Sen yarın oraya gidecek misin?
are you tomorrow there Q-going
'Are you going there tomorrow?'

b. emphatic: Sen yarın mı oraya gidecek sin?
are you Q tomorrow there going
'Are you going there tomorrow?'

c. subj.emph.: Sen mi yarın oraya gidecek sin?
are Q you tomorrow there going
'Are you going there tomorrow?'

Above it has been established on the basis of semantic and syntactic evidence that sentences like (131) and (132) below are derived from the same complex underlying form which is close to (131):

(131) Frantz is the one who helped us.
(132) Frantz helped us.

The question which has not yet been answered is: how do we get from (131) to (132)? As (132) contains much fewer elements than (131), we have to assume that certain elements are deleted, namely the neutral head, the relative clause marker and the copula. Even if all of these deletions occur it does not necessarily imply that the resulting sentence is a simplex. An example of this is TSWANA, a BANTU language. In TSWANA, an SVO language, the position of emphasis is sentence final which corresponds to the position of the focus constituent in a clefted sentence.

(133) a. neutral: rrê orêkilê dikgômo maabane
my-rather he-bourht cattle yesterday
'My father bought cattle yesterday.'

b. subj.emph.: orêkilê dikgômo maabane, rrê
he-bought cattle yesterday, my father
'It is my father who bought cattle yesterday.'

Even though the neutral simplex sentence, e.g. (133.a) differs from the emphatic one only by its word order and by the pause before the emphasized constituent, we cannot consider (133.b) a simplex. The reason is that if (133.b) was a simplex with emphatic word order, we would not be able to account for the pause which is not present in (133.a). The pause signals that (133.b) is an equational sentence.

While in a language like TSWANA the presence or absence of a pause reveals the sentence, it is case assignment that helps us to determine in languages like ENGLISH and GERMAN whether a given sentence is an equational one or not. This is also true for an EAST AFRICAN language like MURLE where an emphatic subject, if part of a cleft, is in the absolute case, while the same emphatic subject is in the nominative if in a non-equational sentence.

(134) a. alaan ak tu-a
chief (abs.) goes back
'The chief is the one who goes back.'

b. ak tu-a alaan-i
goes back chief (nom.)
'The chief goes back.'

In SOMALI, on the other hand, the situation seems to be quite different. As mentioned earlier in SOMALI the form and number of the verb plus the case assignment pattern signal the existence of an underlying cleft construction. However, while in languages like ENGLISH, GERMAN, and MURLE there exist case alternations which can be explained if we assume the reduction of a cleft sentence to a non-equational one, this case alternation does not occur in SOMALI, even though there are other signs that reduction to a simplex has taken place. Consider the following: in SOMALI, an SOV language, extraposition appears to be obligatory in a cleft construction, converting a basic S F Cop order to a F Cop S order, e.g.:

(135) nimán-kíí bàà hɨ́lɨ ʕun-ay?
men-acc. EM meat(acc.) ate (sgl.)
'The men are the ones who ate the meat.'

An alternative form of (135) is (136) below:

(136) hɨ́lɨb nimán-kíí bàà ʕun-ay?
meat(acc.) men-acc. EM ate (sgl.)
'The men ate the meat.'

The only difference between (135) and (136) is that in (135) the emphasized constituent is sentence initial, in (136) it is not.[29] The singular marking on ʕun-ay? 'ate' signals an equational structure with a relative clause as subject. However, while the order of constituents in (135) is that of a cleft structure, the order in (136) does not allow for a cleft interpretation as the emphatic constituent is surrounded by the constituents of the relative clause, which contains the old information. It appears, therefore, that (136) is a simplex, but that in SOMALI, contrary to ENGLISH, GERMAN, and MURLE, case assignment and subject-verb agreement do not apply even if reduction to a simplex has taken place. That is, if we assume that the focus constituent of an equational structure is in the unmarked case (nominative in GERMAN and ENGLISH, accusative or absolute case in MURLE and SOMALI), and if we assume that in a relative clause in SOMALI no subject-verb agreement takes place and the verb therefore is in the unmarked finite verb form (third person singular) then we can say that SOMALI marks an emphatic subject by blocking case assignment and subject-verb agreement in any kind of emphatic structure, whether cleft or simplex. The above analysis of SOMALI is not as ad hoc as it may seem. Kachru (1968) shows that in ergative constructions in HINDI the verb remains in the unmarked masculine singular form, i.e. agreement does not take place, thus

---

[29] Notice that the focus constituent in (136) is in the typical position of emphasis for SOV languages, namely before the verb. For a discussion see below.

signaling certain underlying relations. In SOMALI, absence of case assignment and subject-verb agreement appears to signal that the surface subject is not identical with the underlying subject.[30]

Somali is not the only language in which we can observe that features of the underlying equational structure are preserved in a non-equational surface form. For instance, it was mentioned earlier that KIHUNG'AN uses a special form of negation in relative clauses and in reduced emphatic constructions. That is, the features of the relative clause are retained even if the relative clause as such does not exist any longer:

(137) a. kiim ki a-khoon-in Kipes ku-suum zoon kwe kit
thing pro PA-fail-past Kipes to-buy yesterday is chair
'What Kipes didn't buy yesterday is a chair.'

b. Kipes ka-khoon-in ku-suum kit zoono
Kipes PA-fail-past to-buy chair yesterday
'Kipes didn't buy a chair yesterday.'

The order of constituents in (137.b) indicates that the complex underlying form has been reduced to a simplex, but the features of the underlying cleft structure are retained as in SOMALI. While in the languages discussed so far either the order of constituents or the presence of a pause or a difference in case assignment indicate the status of the sentence, this is not true for some emphatic constructions in RUSSIAN. In RUSSIAN, an SVO language, the focus constituent is in final position in the basic cleft order, e.g. (138):

(138) Tot, kto priexal, eto Vanja.
that who arrived, this Vanya.
'The one who arrived is Vanya.'

Two other emphatic versions of (138) are (139) and (140) below. (139) appears to retain the cleft order with the focus constituent in final position, while (140) has a neutral word order in which the focus constituent is solely marked by loudness.

(139) Priexal Vanja.
arrived Vanya
'Vanya arrived.'

(140) Vanja priexal.
'Vanya arrived.'

---

[30] It seems that SOMALI is one of those languages which marks certain underlying relations by the non-application of otherwise obligatory agreement rules. For a discussion of additional SOMALI data and of the marking hypothesis see Delisle, 1973.

The question is: is (139) a simplex or is it a cleft structure in which all 'superfluous' constituents are deleted. Case assignment will not help us, as Vanja is in the nominative in both (138) and (139). The only indication that (139) is a simplex is the sentence intonation and the absence of a break between the verb and the subject. That (139) is a non-equational structure is also supported by the fact that if the focus constituent corresponds to the object of the relative clause, it is in the nominative in the cleft construction but in the accusative in the reduced construction.

It appears, therefore, that the process of reduction from a complex cleft construction to a surface simplex involves three steps:

1) deletion of all non-essential constituents like the neutral head noun, the relative clause linker, and the copula;
2) reduction to a non-equational string and pruning of superfluous structure;
3) reordering of constituents.

Step 1) by itself does not bring about the formation of a surface simplex as we have seen in the example from TSWANA. Another example is TELUGU which, similar to TSWAÑA, requires an intonation break between the known information and the focus constituent in final position. For the TELUGU data, see examples (125.a) and (125.b) in section 2. The mechanism involved in step 2) can be characterized as predicate lowering in which the predicate is lowered into the relative clause. The question is: what position is the predicate lowered into? It appears that there are two reasonable alternatives:

a. The predicate is lowered into the neutral position which is appropriate to the function the focus constituent will fulfill in the reduced sentence. For instance, if the focus constituent corresponds to the subject of the simplex, it will be placed into initial position in an SVO language, and into medial position in a VSO language. After predicate lowering has taken place, reordering rules apply to generate marked emphatic word orders.

b. The predicate is lowered retaining the same position it had before lowering took place, i.e. the object position. Subsequent reordering generates alternative emphatic orders.

Hypothesis a. claims that the basic order of a reduced emphatic structure is identical to the neutral order of a sentence, while hypothesis b. claims that in the basic order of a reduced emphatic structure the focus constituent is in the position of the object, independent of the function of the focus constituent in the simplex. The difference between hypothesis a. and hypothesis b. can be illustrated with the RUSSIAN example.

Hypothesis a. claims that a cleft structure like (138) is converted into (140) by deletions and predicate lowering and subsequently converted to (139) by a reordering rule. Hypothesis b. claims that (138) is converted to (139) by deletions and predicate lowering and then reordered to (140).

Given the two alternatives, we now have to look for evidence that will substantiate one or the other of these hypotheses. Before we can do that, however, we have to look at the data that we have to account for. That is, what kinds of emphatic word orders occur in simplex sentences. A first look convinces us that it is much more varied and complex than word order variations in cleft sentences. Though initial position is favored by many languages, there also exist final position for subject emphasis in RUSSIAN, SWAHILI and some CHADIC languages like DERA, the second position, after the verb, in MALAGASY (an VOS language) and BADE (an SVO language), and the position before the verb in SOV languages and in ZAPOTEC, an SVO language.

In all the following variations emerge. In SOV languages the contrastive emphasis of a subject results in either OSV or SVO surface orders. Object emphasis is expressed by an OVS surface order. Examples of the above emphasis patterns are found in HUNGARIAN.

(141) HUNGARIAN:

a. neutral SOV: Péter a levelet írja meg.
Peter the letter is writing
'Peter is writing the letter.'

b. subj.emph. OSV: A levelet Péter irja meg.
the letter Peter is writing
'Peter is writing the letter.'

c. subj.emph. SOV: Péter írja meg a levelet.
'Peter is writing the letter.'

d. obj. emph. OVS: A levelet írja meg Péter.
the letter is writing Peter
'Peter is writing the letter.'

In SVO languages object emphasis is expressed by an OSV order in ENGLISH and RUSSIAN and an OVS order in GERMAN. RUSSIAN also allows SOV orders.

(142) a. neutral SVO: I sent the letter yesterday.
b. obj. emph. OSV: The letter I sent yesterday.

(143) GERMAN:

a. neutral SVO: Ich habe den Brief gestern abgeschickt.
I have the letter yesterday sent
'I sent the letter yesterday.'

b. obj. emph. OVS: Den Brief have ich gestern abgeschickt.
the letter have I yesterday sent
'The letter I sent yesterday.'

(144) RUSSIAN:

a. neutral SVO: Boris vzjal knigu.
Boris took book
'Boris took the book.'

b. obj. emph. SOV Boris knigu vzjal.
Boris book took
'Boris took the book.'

c. obj. emph. OSV: knigu Boris vzjal.
book Boris took
'Boris took the book.'

In both GERMAN and ENGLISH there is no marked word order to express subject emphasis, while RUSSIAN and SWAHILI and some CHADIC languages have an VOS order or an VSO order (BADE-CHADIC).

(145) SWAHILI:

a. neutral SVO: manafunzi wadoga wanafungua vitabu.
students small open books
'The small students open the books.'

b. subj. emph. VOS: wanafungua vitabu wanafunzi wadogo.
open books students small
'The small students open the books.'

In VSO languages the emphasized constituent is generally found in sentence initial position. Thus, subject emphasis is expressed by an SVO order while object emphasis results in an OVS order. In ZAPOTEC[31] we can get the following variations:

(146) ZAPOTEC:

a. neutral VSO: gudo xwain biza.
ate Juan beans
'Juan ate beans.'

b. subj. emph. SVO: xwain we ldua7 ati abel
Juan went Oaxaca not Abel.
'Juan went to Oaxaca and not Abel.'

c. obj. emph. OVS: biza gudo xwain ati bel
beans ate Juan not meat
'Juan ate beans and not meat.'

---

[31] The ZAPOTEC data was provided by H. Rosenbaum in a talk given during the California Linguistic Conference, 1973. The title of the talk was: "Conditions on extraction rules: the case from Zapotec."

(148) MALAGASY:

a. neutral VOS: nividy vary ny vehivavy
bought rive the woman
'The woman bought rice.'

b. subj.emph. VSO: nividy ny vehivavy vary
bought the woman rice
'The woman bought rice.'

Summarizing[32] we can make the following statements which in some cases are incomplete due to lack of informations.[33]

		Emphatic variations	
	Neutral	Non-neutral word order	Phonological means only
(149)	S O V	a. O S V b. S V O c. O V S	d. S O V e. S O V
(150)	S V O	a. V O S b. V S O c. O S V d. O V S e. S O V	f. S V O g. S V O
(151)	V S O	a. S V O b. V O S c. O V S d. V O S	e. V S O f. V S O
(152)	V O S	a. V S O	b. V O S c. V O S

If we want to account for all of the emphatic word orders listed above by assuming hypothesis a., where predicate lowering places the focus constituent into the neutral position, we will need many reordering rules which have in common nothing but the fact that they reorder constituents. There will be no way to predict in which direction reordering will move the emphasized constituent. Hypothesis a., therefore implies that reordering for emphatic purposes in reduced emphatic sentences differs from the reordering processes in cleft sentences, as in the latter

---

[32] Even though this section is mainly concerned with marked word orders, sequences in which the neutral word order is retained and emphasis is solely marked by phonological means were included in the above list as these sequences often round off the picture.

[33] For instance I do not know what word order marks object emphasis in MALAGASY, a VOS language.

the whole purpose of reordering is to bring the focus constituent closer to the sentence initial position. On the other hand, if we assume hypothesis b., where predicate lowering places the focus constituent into the object position, we can establish a parallel between cleft sentences and reduced emphatic sentences. In both sentence types the basic position for the focused or emphasized constituent is that position which is filled by the object in a neutral sentence. Furihermore, the only reordering rule that applies to both cleft sentences and their reduced counterparts is the extraposition rule which -- by moving constituents that contain presuppositional material to the end of the sentence -- places the emphasized constituent closer to the beginning of the sentence.

The assumption that the object position is the basic position of emphasis allows us to account for a number of facts for which we have otherwise no explanation. For instance, final position of an emphasized subject is only found in SVO and in VSO languages, e.g. in RUSSIAN and DERA, and in ARABIC and MURLE respectively, but not in SOV and VOS languages. Given hypothesis b., this fact can be explained since the basic position of emphasis in SVO and VSO languages is the final position, i.e. the position of the object, while in SOV and VOS languages, the basic position of emphasis is the one before and after the verb respectively. Thus, hypothesis b. actually predicts that VOS orders should occur in some SVO and VSO languages. That VOS is not found as an emphatic order in all SVO and VSO languages is due to language-specific constraints, which do not permit final position of a subject. Another fact which is explained by hypothesis b. involves SOV languages. Dezső (1968) observed that there exist a hierarchy in SOV languages among the possible emphatic word order alternations. The most basic order, or the order highest in the hierarchy, is the one in which the emphasized constituent is immediately to the left of the verb, e.g. OSV. Only if a language has this order can it also have an additional SVO order. And the highest ranking OSV order corresponds to the basic emphatic order proposed in hypothesis b., as in OSV the emphasized subject is in the position of the neutral object, i.e. the position before the verb. Thus, TURKIC languages in general have OSV orders as the only emphatic order. The same holds true for GALLA, a CUSHITIC language, while HUNGARIAN also allows an SVO order, (see example (141) above).

(154) TURKISH:

a. neutral SOV: Kim sana bunu söyledi?
Who you this told
'Who told you this?'

b. subj.emph. OSV: Sana bunu kim söyledi?
you this who told?
'Who told you this?'

(154) GALLA:

a. neutral SOV: niití-n ilm sāl-tè
woman son she-bore
'The woman bore a son.'

b. subj.emph. OSV: ilm niití-n sāl-tè
son woman she-bore
'The woman bore a son.'

In MALAGASY the position of emphasis is after the verb, i.e. VSO, (see example (148) above). Apparently, no other variations are possible. Given hypothesis a., we will need a special rule for MALAGASY which says: convert a VOS structure to a VSO structure. However, given hypothesis b., the MALAGASY data becomes perfectly regular. Since MALAGASY is a VOS language, the basic position of emphasis is the position after the verb.

It has been shown above that hypothesis b. has to be preferred over hypothesis a. because it simplifies [34] the description and establishes a parallelism between cleft sentences and their simplex counterparts, and because it explain certain data involving specific language types for which hypothesis a. cannot account. Given hypothesis b., we arrive at the following set of basic emphatic positions.

(155) Language type	Basic emphatic positions: Subject emphasis	Object emphasis
S V O	V O S	S V O
S O V	O S V	S O V
V S O	V O S	V S O
V O S	V S O	V O S

If we compare the above list with that of (149) to (152) above which contains all or at least most of the variations found in languages, it is clear that many marked word orders do not correspond to the basic emphatic positions for subjects and objects. Below I will therefore investigate how we can account for the remaining forms. A consequence of assuming that there exist one basic word order for emphatic simplexes is that all other word orders are derived from the basic one by reordering. This reordering takes place by movement rules. During the discussion of alternate word orders in cleft sentences it was argued that the only reordering rule that applies to cleft sentences is an extraposition rule that moves constituents which contain old information to the right of the focus constituent. It was furthermore argued that this rule can apply

---

[34] The fact that hypothesis b. allows for a simpler description will become especially clear later on where it will be shown that only one reordering rule is needed if we assume hypothesis b.

either to the sentence initial constituent or to the pre-focus one. Given that extraposition functions to place the focus constituent into a position of prominence at the beginning of the sentence, there is no reason to expect that the application of this rule is restricted to complex sentences. As a matter of fact, by assuming that extraposition applies to simplexes as well, we can generate all of the marked word orders listed in (149) to (152) as the following derivations show.

(156) Language type: S O V
- a. Subject emphasis:
  - 1. basic emphatic order: O S V (TURKISH, HUNGARIAN, GALLA)
  - 2. extraposition: S V O (HUNGARIAN, VOGUL)
- b. Object emphasis:
  - 1. basic emphatic order: S O V
  - 2. extraposition: O V S (HUNGARIAN)

(157) Language type: S V O
- a. Subject emphasis:
  - 1. basic emphatic order: V O S (RUSSIAN, BANTU, CHADIC)
    extraposition S-initial: O S V
    S V O (GERMAN)
  - 2. basic emphatic order: V O S
    extraposition pre-foc.: V S O (BADE)
    S O V (BULGARIAN, FINNISH)
- b. Object emphasis:
  - 1. basic emphatic order: S V O
    extraposition S-initial: V O S (BULGARIAN)
    extraposition: O S V (ENGLISH, FINNISH)
  - 2. basic emphatic order: S V O
    extraposition pre-foc.: S O V (RUSSIAN)
    extraposition: O V S (GERMAN, SWAHILI, BULGARIAN)

(158) Language type: V S O
- a. Subject emphasis:
  - 1. basic emphatic order: V O S (MURLE, ARABIC)
    extraposition S-initial: O S V
    extraposition: S V O (ARABIC, ZAPOTEC)
  - 2. basic emphatic order: V O S
    extraposition pre-foc.: V S O (ARABIC)
    extraposition: S V O (ARABIC, ZAPOTEC)

b. Object emphasis:
   1. basic emphatic order: V S O
      extraposition S-initial: S O V (ZAPOTEC)
      extraposition: O V S (ZAPOTEC, ARABIC)
   2. basic emphatic order: V S O
      extraposition pre-foc.: V O S
      extraposition: O S V

(159) Language type: V O S
   a. Subject emphasis:
      1. basic emphatic order: V S O (MALAGASY)
         extraposition: S O V
   b. Object emphasis:
      1. basic emphatic order: V O S
         extraposition: O S V

The above list shows that in some instances the emphatic word order corresponds to the neutral one.[35] Given hypothesis b. this is always the case for object emphasis since hypothesis b. claims that the neutral word order and the basic position of emphasis for objects are identical. However, this is not true for subject emphasis. Given the basic emphatic order and extraposition, only SVO and VSO languages should allow an SVO and an VSO order respectively. Examples of these orders are found in ENGLISH and ARABIC.

(160) Bill didn't see John, but Mary saw him.

(161) ARABIC:
kataba Ali ar-risāla
wrote Ali the-letter
'Ali wrote the letter.'

In SOV and VOS languages, on the other hand, no SOV and VOS order is generated by the rules proposed so far. Consequently, if there are SOV and VOS languages that do have an SOV or a VOS order respectively, the rule component has to be modified to the effect that in addition to extraposition there exists a regrouping process which generates the neutral word order, converting for example an OSV order into an SOV one. Evidence from HUNGARIAN indicates that such a regrouping process which generates neutral word orders does not exist, as in HUNGARIAN the neutral word order cannot be used to express subject emphasis, e.g. (163.b):

---

[35] It appears that identity between the neutral word order and an emphatic one is only possible if the emphatic constituent is marked by some other means, for instance by phonological ones.

(162) neutral: János könyvet vett.
John book bought
'John bought a book.'

(163) subj. emph.: a. János vett egy könyvet.
'John bought a book.'

b. *János könyvet vett.

The ungrammaticality of (163.b) indicates that the neutral order is an alternative emphatic order only if it happens to be one of the orders generated by predicate lowering (e.g. an SVO order in ENGLISH), and by the extraposition rule (e.g. an SVO order in GERMAN), but not if it has to be generated by a special ordering convention that establishes the neutral order (e.g. an SOV order in HUNGARIAN).

During the discussion of the extraposition rule in cleft sentences it was mentioned that data from marked word orders in non-equational sentences supports the claim that extraposition and not focus fronting is the rule that is responsible for the regrouping of constituents. The failing of the fronting rule as a possible alternative to extraposition was not that obvious in clefted sentences, mainly because a cleft sentence can be more easily divided into two parts, the subject NP and the predicate, which may or may not contain the copula. But as soon as more than two elements are involved, fronting can be shown to be totally inadequate. A fronting rule predicts: 1) that the focus constituent is placed into sentence-initial position, and 2) that the order of the remaining constituents is not affected. Both of these predictions are violated by many of the above examples. Consequently, if we assume the fronting rule, we will need a number of highly idiosynchratic and ad hoc rules which would have to apply after the fronting rule to generate marked surface orders. Let us look at a few examples. In RUSSIAN object emphasis can result in an SOV order. To derive this order from an underlying SVO order by fronting, we have two alternatives: 1) front the object and have an additional rule that moves the object back before the verb, 2) have an alternative rule for RUSSIAN that places an emphatic object before the verb. In ZAPOTEC, a VSO language, SOV signals object emphasis. The reader can see for himself that a conversion of VSO to SOV by fronting would lead over a very tortuous road. BADE, a CUSHITIC SVO language, would again require a very specific rule to generate VSO subject emphasis, namely a rule that places an emphatic subject behind the verb. Extraposition, on the other hand, can generate all of the attested marked emphatic word orders without the need for any additional rules. All that is required are language-specific constraints on the rule that state which orders are possible in a given language. The fact that I did not find examples for some of the predicted word orders is not

especially disturbing. The most likely reason is that not enough languages were investigated,[36] but it is also possible that certain orders will never occur because of the existence of a still undiscovered principle.

Summarizing this section on predicate lowering and marked word orders we can say that there is ample evidence in support of a rule that reduces complex cleft sentences to surface simplexes, and that there is furthermore considerable evidence that during this reduction the focus constituent is placed into the position which is held by the object in a neutral statement. This position I have called the basic emphatic position. Subsequent to the reduction and placing of the focus constituent extraposition can apply with the effect that the focus constituent moves closer to the front of the sentence. Whether or not the focus constituent will be moved all the way to the front or only "half-way" depends on language-specific constraints. Also language-specific is whether the extraposition rule will apply to the sentence-initial constituent, or to the pre-focus one. Pre-focus application is only possible in those languages in which the basic position of emphasis is clause final, i.e. in SVO and VSO languages. This accounts for the fact that SVO and VSO languages have more alternative word orders than SOV and VOS languages, in which the basic emphatic position is medial.

The question which has not been mentioned at all in this paper and which will have to be investigated thoroughly is: what happens if more than the basic subject, object, verb constituents are present? Does extraposition influence the order of those additional constituents and if yes, how? Or are those constituents left outside of the regrouping process? The very limited sample from GERMAN and ENGLISH I looked at suggests the latter, but considerable research in this area is still required before we can come to any definite conclusions.

6. Phonological manifestations of emphasis

Above it was briefly mentioned that if the marked word order is identical with the neutral one, heavy stress on the emphasized constituent is often used to mark the emphasized constituent. However, emphatic or contrastive stress is not only used in the above cases, but can also occur on the emphasized constituent in cleft sentences and in non-neutral word orders, depending on how much a speaker wants to

---

[36] One of the greatest problems in working with grammars is that if they discuss word order at all, they hardly ever mention which of the constituents in a marked word order is the emphasized one. Thus, if it is said that an SVO languages also has an alternative OSV order, it can mean either OSV or OSV, or possibly even OSV.

underline a certain contrast. Phonological manifestations of emphasis appear to be secondary, that is, all languages that use increased loudness or a tonal change to indicate emphasis also have syntactic means like cleft construction and marked word orders to express emphasis. Though most of the language I investigated can make use of contrastive stress -- INDO-EUROPEAN languages, FINNO-UGRIC languages, most BANTU languages, most SEMITIC languages, MALAYO-POLYNESIAN languages, and CHINESE -- some languages do not use it. For instance, VEI, a MANDAN language of West Africa and AMHARIC do not use loudness or any other phonological mechanism to signal emphasis. In KIHUNG'AN, on the other hand, the emphasized constituent undergoes a change in tone, i.e. a low tone becomes a high one.[37]

(164) neutral: Kìpès ká-swiim-ín kìt zɔ̀ɔ̀n
Kipes buy-past chair yesterday
'Kipes bought a chair yesterday.'

(165) obj. emph.: Kìpès ká-swíím-in kít zóónó
Kipes buy-past chair yesterday
'Kipes bought a chair yesterday.'

(166) subj. emph.: Kìpés ká-swíím-ín kìt zóónó
Kipes buy-past chair yesterday
'Kipes bought a chair yesterday.'

While supra-segmentals play an important role in marking the emphatic character of a constituent, other phonological alternations were not observed. That is, no language was found in which phonological segments were systematically converted into other phonological segments to signal emphatic contrast, though increased loudness is usually coupled with a slowing down of the speed of the utterance which in turn results in a lengthening of the stressed vowel.

It would go beyond the framework of this paper to investigate in detail where and how contrastive stress is assigned. I will therefore only make a few remarks which might lead to further investigations. As the term already indicates, contrastive emphasis contrasts, implicitly or explicitly, two or more elements which are members of parallel sentences in the deep structure. Contrastive emphasis is assigned on the basis of syntactic parallelism, that is in two parallel sentences the non-identical constituents are stressed. One of the parallel structures can be implied as in (167):

---

[37] From the limited data available it appears that if an emphasized word has more than one low tone, it is the last tone that will be changed.

(167) a. Bill helped to repair the car.
b. No, it was John who helped to repair the car.

In (167.b) a full sentence is deleted except for the no, namely the sentence:[38]

(167) c. No, it was not Bill who helped to repair the car.

While (167.c) is deleted anaphorically with (167.a) as the antecedent, deletion of a parallel sentence can also take place if the contrasted constituent in that sentence is left vague as in (168.a).

(168) a. No, it was not John who helped, it was somebody else.
b. No, it was not John who helped.

(168.a) and (168.b) are semantically identical, i.e. no information is lost by the deletion.

Given that the parallel structure can be destroyed by deletion of one of its members, and assuming that contrastive stress is assigned on the basis of this parallel structure, it appears that stress has to be assigned before deletion takes place.[39] Both members of the parallel structure have to be present in some languages if the first member is negated.

(169) GERMAN:
a. Nicht Jürgen ist gekommen, sondern Hans.
'Jürgen didn't come, but Hans.'
b. ?Nicht Jürgen ist gekommen.

(170) HUNGARIAN:
a. Nem Janos repült New York-ba, hanem Imre.
Not John flew New York-to, but Imre
'John didn't fly to New York, but Imre.'
b. ?Nem János New York-ba repült.

7. Schachter's promotion analysis

The analysis on which this paper has been based is generally called the "pseudo-cleft analysis", as it assumes that the underlying form of cleft sentences is a so-called pseudo-cleft as below:

(171) The one who left early was Frank.

---

[38] For further discussion and evidence for the existence of (167) in the underlying form of (167.b), see Harries, 1972.

[39] For a discussion as to where stress is assigned, see Bresnan, 1971 and 1972, and Berman & Szamosi, 1972. Bresnan claims that "intonation depends systematically on underlying syntactic structure" (1972: 326).

The pseudo-cleft analysis was rejected by Schachter (1973) for the following reason: there exists a set of cleft sentences for which there exists no corresponding set of pseudo-clefts. For instance, if we want to claim that all cleft sentences are derived from underlying pseudo-clefts, then we have to assume that (172.d) below is derived from an ungrammatical form, namely (172.f).

(172) a. John bought the book for Bill.
b. It was Bill that John bought the book for.
c. It was Bill for whom John bought the book.
d. It was for Bill that John bought the book.
e. The one that John bought the book for was Bill.
f. *The one that John bought the book was for Bill.

By rejecting the pseudo-cleft analysis on the basis of the ungrammaticality of (172.f), Schachter takes the position that a grammatical surface form should not be generated from an ungrammatical underlying form. However, the same problem is encountered with some adjectives which can occur in attributive but not in a predicate position. That is, for these adjectives relative clause reduction and front moving is obligatory.

(173) a. *My hand which is left is hurt.
b. My left hand is hurt.

But even if we reject the derivation of a grammatical surface form from an ungrammatical underlying form, this does not necessarily imply that clefts are not derived from pseudo-clefts. There are certain properties about sentences like (172.d) which show that they are not equivalent to sentences like (172.c). The most striking difference is that the that in (172.d) is not a relative pronoun as in (172.b), but rather a complementizer. That this is so is more obvious in GERMAN and FRENCH where the complementizer and the relative pronoun have different forms. In the sentences below, the a. versions contain a relative pronoun, while the b. versions contain a complementizer.

(174) a. Das ist Hans, für den ich das Buch gekauft habe.
'It is Hans for whom I bought the book.'
b. Das ist für Hans $\left\{\begin{matrix}\text{daß}\\ \text{*für den}\\ \text{*dem}\end{matrix}\right\}$ ich das Buch gekauft habe.
'It is for Hans that I bought the book.'

(175) a. C'est Jean pour qui j'ai acheté le livre.
'It is John for whom I bought the book.'
b. C'est pour Jean $\left\{\begin{matrix}\text{que}\\ \text{*pour qui}\\ \text{*qui}\end{matrix}\right\}$ j'ai acheté le livre.
'It is for John that I bought the book.'

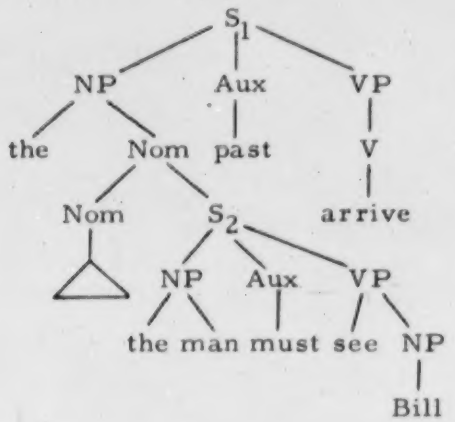
And even in ENGLISH, that in (2.d) cannot be replaced by a wh-pronoun.

(172) g. *It was for Bill $\left\{\begin{matrix}\text{for whom}\\ \text{whom}\\ \text{who}\end{matrix}\right\}$ John bought the book.

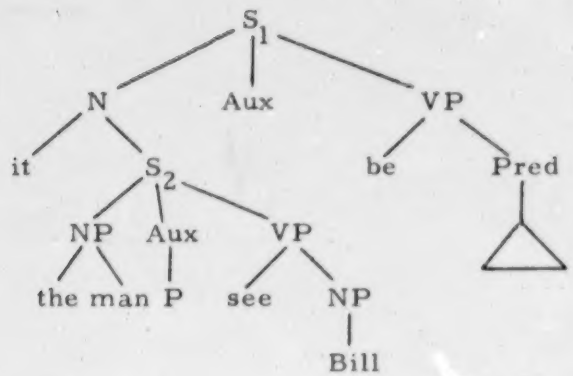
The above data suggests that sentences like (172.c) have a different underlying form from sentences like (172.d) and the latter sentence type therefore does not necessarily constitute counter-evidence to an analysis which claims that all so-called pseudo-clefts underly all other cleft constructions.

However, Schachter considers sentences like (172.d) counter-evidence and therefore proposes the promotion theory. This theory claims that we can account for the fact that relative clauses and cleft sentences show so many similarities on the basis that both undergo the same rule, namely the promotion rule. This rule copies noun phrases out of an embedded S onto a dummy △ in a higher S. That is, relative clauses have an underlying form as in (176) while cleft sentences have an underlying form as in (177).

(176)



(177)



Let us first look at the rule or rules needed to replace the dummy symbols with a constituent. In order to convert (176) into a surface form, one[40] of the NP's of S_2 has to be copied onto the dummy symbol to lexicalize the head of the relative clause. In (177), on the other hand, a noun phrase is copied onto the dummy node in the predicate. Thus, the structural descriptions of the two rules are quite different. What they have in common is that they copy a constituent of a lower S onto a dummy symbol in a higher S. Another difference between the two rules is the kind of constituents they can raise. While the head of a relative clause can only be a noun phrase, the predicate of a cleft sentence is much less restricted. It can be a noun phrase, an adjective, or an adverb. Furthermore, the noun phrases that are raisable are more restricted in a relative clause than in a cleft sentence. For instance, if we want to derive surface form S_A like (172.d) from an underlying form like (179), the promotion rule has to copy the noun phrase and the preposition, and subsequent rules that effect the original constituent in the embedded S have to be made sensitive to whether a noun phrase was copied with a preposition or without it. If it was copied without a preposition, the original constituent is relativized (e.g. (172.b) and (172.c)). However, if the original constituent was copied with a preposition, then the original has to be deleted and a that-complementizer has to be introduced (e.g. (172.d). Consequently, even in Schachter's framework the derivation of a sentence like (172.d) is quite different from the derivation of a sentence like (172.b) and requires special constraints. Besides the constraints listed above, Schachter also has to stipulate that the rule of extraposition from it is obligatory, even though this kind of rule is usually considered optional.

Another drawback of the promotion hypothesis is that it cannot relate clefts and pseudo-clefts, that is it makes the claim that sentences like (178) and (179) below are completely unrelated.[41]

---

[40] Schachter himself points out that the fact that the promotion rule can choose any one of the noun phrases in the embedded S violates the assumption held by most linguist today, namely that the underlying syntactic structure of a sentence must have a unique semantic interpretation. Schachter's proposal therefore requires a change in the theory to the effect that the semantic interpretation of a sentence takes place somewhere between the deep structure and the surface structure representation. A similar change in the theory is needed to accomodate cooccurrence constraints.

[41] Schachter does not attempt to relate (178) or (179) above to the emphatic simplex, i.e. 'John arrived.'

(178) The one who arrived is John.

(179) It is John who arrived.

Given that (178) and (179) are unrelated in Schachter's hypothesis, he also fails to account for the facts discussed by Akmajian, which support the hypothesis that clefts are derived from pseudo-clefts. The data presented by Akmajian involves the occurrence of certain pronouns in the relative clause of the cleft sentence. These pronouns cannot be accounted for if we assume an underlying it-cleft, but are perfectly regular if we assume an underlying pseudo-cleft with a third person head. There exist similar evidence in GERMAN involving the relative pronoun. Consider the following data:

(180) a. Es ist bloßer Neid, was aus ihm spricht.
It is pure envy, what out him speaks
'It is pure envy, what motivates him.'

The relative pronoun was [neuter] cannot have Neid [masculine] 'envy' as antecedent, as GERMAN has obligatory gender agreement between the head and the relative pronoun. However, if we assume that (180.a) is derived from (180.b), then we can account for the occurrence of was in (180.a).

(180) b. Dasjenige (Ding), was aus ihm spricht, ist bloßer Neid.
That (thing), what out ihm speaks, is pure envy
'That (thing), what motivates him, is pure envy.'

Further evidence that was cannot have Neid as antecedent is that was as a relative pronoun only occurs in certain environments, e.g. if the head has been deleted or is somehow undetermined or vague. Otherwise the neuter relative pronoun is das 'that'.

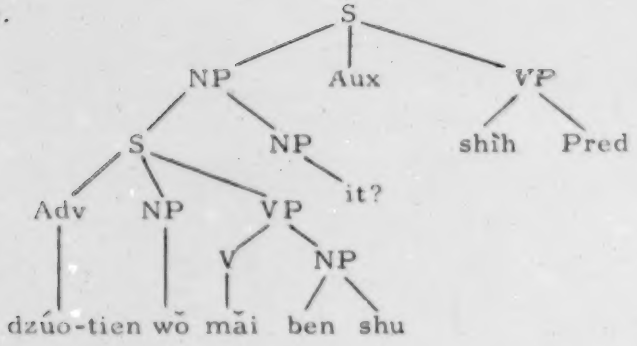
(181) Alles was er tut ist sinnlos.
'All what he does is senseless.'

Finally, I will discuss a more general problem that Schachter's analysis faces. It has been shown by Sanders and Tai (1972) that in languages like MANDARIN CHINESE, BENGALI, SYRIAN ARABIC, etc., an object cannot be removed from the VP, i.e. it cannot be deleted or topicalized.[42] This constraint would have to be violated by Schachter's analysis as (182.a) below would be generated from an underlying form like (182.b), where the future focus constituent shu 'book' is the object of the relative clause.

(182) a. dzúo-tien wǒ mǎi de shìh běn shu
yesterday I buy RM is clas. book
'What I bought yesterday is a book.'

[42] For a discussion of this data and its analysis see Sanders and Tai, 1972 and Harries, 1973b.

b.



Summarizing this brief discussion of Schachter's proposal we have to say that as it stands the promotion hypothesis proves to be inadequate in too many respects to be considered a viable alternative to the pseudo-cleft analysis. Schachter's whole argument against the pseudo-cleft analysis is that there exist certain cleft sentences which do not have any pseudo-cleft counterparts in ENGLISH, e.g. example (172.d). It has, however, been shown above that those sentences are not only problematic for the pseudo-cleft analysis, but also for the promotion analysis. Schachter claims that his analysis explains why cleft sentences and relative clause are very similar in many languages. But all this relationship is built on, according to Schachter, is that two different underlying structures undergo a similar, though by no means identical, rule. This appears to be a rather weak link to account for such obvious similarities, especially as the similarities between relative clauses and clefts are not brought about by the promotion rule. In other words, the similarities -- for instance the constraints on the embedded verb -- are not a result of the application of the promotion rule, and consequently the promotion rule cannot be the reason for these similarities. The pseudo-cleft analysis, on the other hand, has been shown to be adequate not only for ENGLISH, but on a universal basis. It accounts for all the varied data presented in this paper and does not require any rules that are not needed elsewhere in the grammar.

8. Conclusion

In conclusion I will summarize the major claims made in this paper and will briefly discuss an area not touched on so far, namely the relation of cleft sentences to word questions and their answers.

8.1 Summary

The basic claim of this paper is that all sentences that contain a contrastively emphasized noun phrase are derived from underlying equational sentences. In these equational sentences the known information is represented by the subject, while the new and focused information is represented by the predicate noun phrase. To the underlying cleft sentence rules

can apply to generate different surface structures. In most languages the full cleft structure is grammatical, but in some languages deletion and/or reordering rules must apply obligatorily. The two sets of rules that account for the variations in the surface form of emphatic structures are deletion rules, i.e. head deletion, copula deletion, and relative marker deletion, and reordering rules, i.e. extraposition, and predicate lowering. It was argued that extraposition of the subject and not fronting of the focus constituent accounts for the reordering of constituents in a cleft sentence. While extraposition does not effect the underlying equational structure, predicate lowering reduces the original equational structure to a surface form which is no longer equational, thus relating sentences like (183.a) and (183.b).

(183) a. The one who fell from his bike is a friend of John's.
b. A friend of John's fell from his bike.

It is argued that the focus constituent is lowered into the unmarked position of emphasis which is the position of the object or complement. This is also true for cleft sentences, i.e. both in an equational and in a non-equational emphatic structure the basic position of emphasis corresponds to the position of the object in a neutral sentence. Only if we make this assumption, can we account in a principled way and without any ad hoc rules and constraints for all of the emphatic word order variations that can occur.

Finally, in a discussion of Schachter's promotion hypothesis it is shown that Schachter's approach not only fails to account for all of the data, but it also requires major adjustments in the theory which are not otherwise justified.

8.2 Word questions and their answers

I have argued elsewhere[43] that all word questions and their answers closely resemble contrastively emphasized structures and are in fact derived from underlying cleft sentences. Functionally, questions and their answers and contrastively emphasized structures are very similar. A contrastively emphasized structure contrasts one or several members of a related set, while a word question requests the identification of one or several members of a set. This identification is provided in an appropriate answer, which names those members of the set which fit the description provided in the question and contrast it with all other members of the set which do not fit the description. For instance, in (184) below, it is requested that the answering person only names that member of the specified set who is sick and contrasts him with all the other members of the set who did not get sick.

---

[43] Harries, 1972.

(184) a. Who flew to the moon and got sick on the way?
b. Frank Miller (but not Bill Johnson, John Smith, etc.)

The same parallelism that exists in contrastively emphasized structures between the equational version and its non-equational counterpart also exists for questions and for their answers. That is, (185) below is a paraphrase of (184.a) in that it contains the same presuppositions.

(185) Who is the one who flew to the moon and got sick on the way?

The hypothesis that emphatic constructions and word questions are related has been put forward before for a number of languages. Takizala (1972) points out that in KIHUNG'AN word questions are subject to the same constraints on negation and object pronoun infixation as relative clauses and emphatic constructions, and he therefore concludes that word questions have an underlying cleft structure. In CHIPPEWA, word questions behave like sentences with embedded relative clauses, that is word questions always contain a subordinate verb form. In CHADIC languages, word questions generally share characteristics with emphatic constructions, i.e. the same constraints hold for both sentence types. While in the above-discussed cases word questions reveal their cleft origin, there are some languages in which this is also true for the answer. For instance, Hutchinson (1969) observes that in TEMNE both a question and its answer have to be clefts, e.g. (186.a) and (186.b). (186.c), the non-cleft counterpart to (186.b), is not an acceptable answer to (186.a).

(186) a. kanɛ n nəŋk a?
'Who was it that you saw?'

b. ɔbại kɔnɛ i nəŋk.
'It was the chief whom I saw.'

c. ɔ i nəŋk ɔbại
'I saw the chief.'

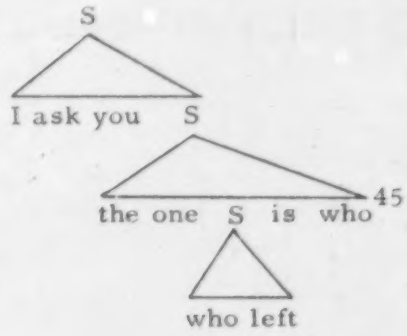
It is claimed, then, that word questions and their appropriate answers have an underlying equational structure in which the focus constituent of the question, i.e. the question word, is only partially specified. This partial specification is filled in by the focus constituent of the answer. The only other difference between a question and its answer is the performative involved. The subject noun phrases are identical as they contain the presuppositions.[44] Thus word questions

---

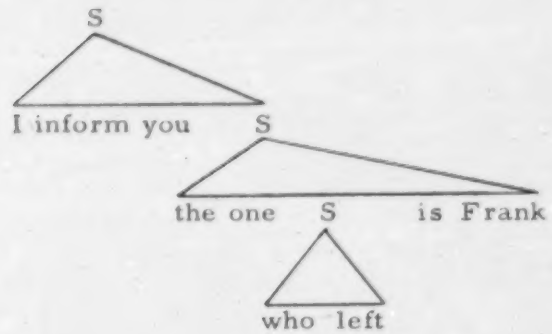
[44] In my thesis I argued that an answer to a given question is appropriate only if the two share the same presuppositions.

have an underlying structure like (187.a), answers like (187.b).

(187) a.



b.



Both word questions and their answers can be the input to the rules discussed in this paper. For instance, in some languages like ENGLISH and GERMAN, extraposition has to apply to (187.a) while in other languages, like the BANTU ones, this is not necessarily so. Thus, the rules that are needed to generate word questions are the same rules that apply to non-interrogative equational structures, and no specific 'question rules' are required.

Given the above hypothesis the next step will be to determine what kinds of language specific constraints are found in word questions. In other words, we will have to investigate what is and what is not a possible constraint on a word question.

---

[45] Truitner, Dunnigan (1972) mention that the question word in CHIPPEWA appears to be complex. There is also some evidence from GERMAN involving negation which suggests that the predicate nominal is part of a complex structure of the form: the one who is Frank. See Harries (1972).

BIBLIOGRAPHY

Akmajian, A. 1970. On deriving cleft sentences from pseudo cleft sentences. Linguistic Inquiry, 1.2. 149-168.

Ashton, E.O. 1944-1970. Swahili grammar. London.

Bach, E. 1967. Have and be in English syntax. Language, 43.2.562-485.

_____ and S. Peters. 1968. Pseudo-cleft sentences, preliminary version. University of Texas, Austin.

Balmer, W.T. and F.C. Grant. 1929. A grammar of the Fante-Akan language. London.

Bell, C.R. 1968. The Somali language. New York.

Benveniste, E. 1960. "Être" et "avoir" dans leurs fonctions linguistiques. BSL, 55. 113-34.

Berman, A. and M. Szamosi. 1972. Observations on sentential stress. Language, 48.2.304-325.

Beythan, H. 1943. Praktische Grammatik der Tamilsprache. Leipzig. Vlg. O. Harrassowitz.

Borras, M. and R.F. Christian. 1961. Russian syntax. Oxford.

Bresnan, J. 1971. Sentence stress and syntactic transformations. Language, 47.2.257-81.

_____. 1972. Stress and syntax: a reply. Language, 48.2.326-342.

Chomsky, N. 1970. Remarks on nominalization. In Readings in English transformational grammar, ed. by R. Jacobs and P. Rosenbaum. Waltham, Mass. 184-221.

Cohen, M. 1970. Traité de langue amharique. Paris.

Cole, D.T. 1955. An introduction to Tswana grammar. London.

Cowel, M.W. 1964. A reference grammar of Syrian Arabic. Arabic Series, No. 7, ed. by R.S. Harrell. Institute of Languages and Linguistics, Georgetown University.

Crazzolara, J.P. 1960. A study of the Logbara (Ma'di) language. Oxford University Press.

Darden, B. 1969. On the question of the copula. In Papers of the Fifth Regional Meeting of the Chicago Linguistic Society. 30-35.

Delisle, G. 1973. Non-standard concord and the marking hypothesis. WPLU, 11. 85-138.

Dezső, L. 1968. Einige typologische Besonderheiten der ungarischen Wortfolge. In Acta Linguistica Academiae Scientiarum Hungaricae Tomus 18 (1-2), 125-159.

Ellis, J. and L. Boadi. 1969. 'To be' in Twi. In The verb 'be' and its synonyms, ed. by J. Verhaar, Part 4, V. 9, 1-65.

Ferguson, C. 1972. Being in Bengali, with a note on Amharic. In The verb 'be' and its synonyms, ed. by J. Verhaar, Part 5, V. 4, 74-114.

Halliday, J. 1967. Grammar, society, and the noun. Inaugural lecture 1966, University College, London.

Harries, H. 1972. Cleft sentences, questions, and presupposition sharing. Doctoral dissertation, University of Minnesota.

_____. 1973a. A note on double negation and modals as main verbs. WPLU, 11. 211-17.

_____. 1973b. Coordination reduction. WPLU, 11. 139-209.

Horne, E.C. 1961. Beginning Javanese. Yale Linguistic Series, 3. New Haven.

Hutchinson, L. 1969. Pronouns and agreement in Temne. Doctoral dissertation, University of Indiana.

Ibañez, T. 1970. Emphase and der Bereich der Negation: Satz - vs Satzgliednegation. In Arbeitspapier, Nr. 9, Institut für Sprachwissenschaft, Universität zu Köln.

Jespersen, O. 1949. A Modern English Grammar, Vol. VII, part 4. Heidelberg. (originally published in 1909).

_____. 1969. Analytic syntax. Transatlantic Series in Linguistics. New York.

Kähler, H. 1965. Grammarik der Bahasa Indonésia. Wiesbaden.

Keenan, E.L. 1972. Relative clause formation in Malagasy (and some related and not so related languages). In Chicago Which Hunt Papers from the Relative Clause Festival, Chicago Linguistics Society, 169-189.

Koelle, S.W. 1884. Outlines of a grammar of the Vei language. London. (1968 edition).

Krishnamurti, B.H. and P. Sivananda Sarma. 1968. A basic course in Modern Telugu. Hyderabad.

Kuno, S. 1972. Functional sentence perspective: a case study from Japanese and English. Linguistic Inquiry, III.3. 269-320.

Lakoff, G. 1965. On the nature of syntactic irregularity. Doctoral dissertation, Indiana University.

Lehtinen, M. 1963. Basic course in Finnish. Uralic and Altaic Series, V.27, Indiana University Publications.

Lees, R.B. 1963. Analysis of the "cleft sentence" in English. In Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung, Bd. 16, Heft 4, 371-88.

Lukas, J. 1937. A study of the Kanuri language. Oxford University Press.

Malzac, R.P. 1960. Grammaire malgache. Paris.

McQuown, N.A. and S. Koylass. 1945. Spoken Turkish. New York.

Meskill, R.H. 1970. A transformational analysis of Turkish syntax. The Hague: Mouton.

Myers, Amy. 1970. A grammar of Tera. University of California Press.

Ross, J.R. 1969. Adjectives as noun phrases. In Modern Studies in English, ed. by D. Reibel and S. Schane. 352-60.

_____. 1969b. Auxiliaries as main verbs. In Studies in Philosophical Linguistics, Series I, ed. by W. Todd. Evanston, Ill.: Great Expectations Book Sellers and Publishers. 77-102.

Sapir, D. 1965. A grammar of Diola-Fogny. West-African Language Monograph, No. 3. Cambridge.

Sanders, G. and J. Tai. 1972. Immediate dominance and identity deletion. Foundations of Language, 8.2. 161-97

Schachter, P. 1973. Focus and relativization. Language, 49.1. 19-46.

Schachter, P. and F.T. Otanes. 1972. Tagalog reference grammar. University of California Press.

Schuh, R.C. 1971. Reconstructions of the syntax of subject emphasis in certain Chadic languages. In Studies in African Linguistics, Supplement 2. 67-78.

Schwartz, A. 1971. General aspects of relative clause formation. WPLU, 6. 139-71.

Takizala, A. 1972. Focus and relativization in Kihung'an. In Studies in African Linguistics, 3.2. 259-87.

Thank Bĩhn, D. 1971. A tagmemic comparison of the structure of English and Vietnamese sentences. The Hague: Mouton.

Thompson, L.C. 1969. A Vietnamese grammar. Seattle: University of Washington Press.

Tompa, J. 1968. Ungarische Grammatik. The Hague: Mouton.

Truitner, D. and T. Dunnigan. 1972. Wh-questions in Ojibwe. In Papers from the 8th Regional Meeting of the Chicago Linguistics Society. 359-67.

Ultan, R. 1969. Some general characteristics of interrogative systems. WPLU, 1. 41-63a.

Vân Chiuh, T. 1970. Structure de la langue vietnamienne. Paris.

Welmers, W.E. 1964. The syntax of emphasis in Kpelle. In The Journal of West African Languages, 1.1. 13-26.

Westermann, D. 1907. Grammatik der Ewe-Sprache. Berlin.